

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2005, 08:29:23 ; Search time 1832 Seconds  
(without alignments)  
1243.120 Million cell updates/sec

Title: US-10-349-143-1

Perfect score: 47

Sequence: 1 ccacagctgattagaat.....gatcacatttcaatttt 47

Scoring table: IDENTITY\_NUC

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hgt:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sbs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46.6	99.1	47	6	AR288266 Sequence
2	46.6	99.1	93153	2	AF257499 Homo sapi
3	46.6	99.1	182151	9	AC011835 Homo sapi
4	46.6	99.1	185463	9	AC016065 Homo sapi
5	46.6	99.1	191377	9	AC018398 Homo sapi
6	46.6	99.1	240825	6	AX087869 Sequence
7	46.6	99.1	240825	6	AX523960 Sequence
8	29.4	62.6	142552	9	HS5188 Human DNA
9	29.4	62.6	160530	2	AC025631 Homo sapi
10	29.4	62.6	266079	2	AL391060 Homo sapi
11	27.6	58.7	40479	9	AL356492 Human DNA
12	27.6	58.7	70766	2	AL356492 Homo sapi
13	27.6	58.7	118335	8	AC006193 Arabidops
14	27.2	57.9	177483	9	AC108046 Homo sapi
15	27.2	57.9	225586	5	AX469934 Homo sapi
16	27.2	57.9	237042	2	BX908795 Zebrafish
17	26.4	56.2	160796	2	AL592222 Danio rer
18	26.4	56.2	169994	5	BX31875 Zebrafish
19	26.4	56.2	197773	2	BX936440 Danio rer

20	26.4	56.2	203451	10	AC023173 Mus muscu
21	26.4	56.2	218449	10	AC122835 Mus muscu
22	26.2	55.7	213417	2	AC119515 Rattus no
23	26.2	55.7	217538	2	AC098034 Rattus no
24	26.2	55.7	242655	2	AC107465 Rattus no
25	26.2	55.7	1313096	2	AC131879 Rattus no
26	26	55.3	138307	10	AL808014 Mouse DNA
27	26	55.3	162058	2	AC135447 Rattus no
28	26	55.3	238527	10	AC129700 Rattus no
29	26	55.3	242422	2	AC107551 Rattus no
30	26	55.3	321552	2	AC109528 Rattus no
31	25.8	54.9	128301	9	AC133480 Homo sapi
32	25.8	54.9	154109	2	CR376803 Danio rer
33	25.8	54.9	155754	2	AC011273 Rattus no
34	25.8	54.9	162087	2	AC011576 Homo sapi
35	25.8	54.9	164423	2	CR361552 Danio rer
36	25.8	54.9	177203	2	AC024531 Homo sapi
37	25.8	54.9	204821	2	BX276111 Danio rer
38	25.8	54.9	225216	10	AL844888 Mouse DNA
39	25.6	54.5	1913	9	BC033135 Homo sapi
40	25.6	54.5	106302	9	AC080011 Homo sapi
41	25.6	54.5	150939	2	AC091676 Homo sapi
42	25.6	54.5	151386	2	AC015731 Homo sapi
43	25.6	54.5	160119	2	AC130400 Homo sapi
44	25.6	54.5	165407	2	AC025136 Homo sapi
45	25.4	54.0	6890	8	AB010274 Schizosac
46	25.4	54.0	23193	8	SPAC926
47	25.4	54.0	165825	2	BX511301 Danio rer
48	25.4	54.0	186858	10	AC130210 Mus muscu
49	25.2	53.6	65294	9	AC110791 Homo sapi
50	25.2	53.6	92587	8	AC096665 Homo sapi
51	25.2	53.6	93375	8	AC139853 Medicago
52	25.2	53.6	93375	8	AC139853 Medicago
53	25.2	53.6	95155	9	AC078950 Homo sapi
54	25.2	53.6	112823	2	AC013301 Homo sapi
55	25.2	53.6	150194	5	BX649510 Zebrafish
56	25.2	53.6	165349	2	AC093665 Homo sapi
57	25.2	53.6	166458	10	AC136455 Mus muscu
58	25.2	53.6	166857	2	AC098589 Homo sapi
59	25.2	53.6	199523	9	AC012014 Homo sapi
60	25.2	53.6	220892	2	AC111605 Rattus no
61	25	53.2	24521	10	BX649344 Mouse DNA
62	25	53.2	42271	2	AC100593 Mus muscu
63	25	53.2	88982	5	AL732538 Zebrafish
64	25	53.2	110000	1	BX897699_06
65	25	53.2	110000	1	BX897699_07
66	25	53.2	110000	2	BX546462_1
67	25	53.2	118016	10	AL929405 Mouse DNA
68	25	53.2	121466	9	AC105310 Homo sapi
69	25	53.2	166409	2	CR774190 Danio rer
70	25	53.2	177949	2	CR513782 Danio rer
71	25	53.2	179475	5	BX001025 Zebrafish
72	25	53.2	180320	10	AC127588 Mus muscu
73	25	53.2	181069	2	CR762471 Homo sapi
74	25	53.2	183230	5	BX649488 Zebrafish
75	25	53.2	184923	2	CR847497 Danio rer
76	25	53.2	186379	2	CR405691 Danio rer
77	25	53.2	186645	10	AL807773 Mouse DNA
78	25	53.2	189787	2	CR293534 Danio rer
79	25	53.2	240548	2	AC094036 Rattus no
80	25	53.2	254657	5	BX004878 Zebrafish
81	25	53.2	328252	2	AC120720 Rattus no
82	25	53.2	328252	2	CR354592 Danio rer
83	25	53.2	47557	2	CR354592 Homo sapi
84	24.8	52.8	66610	9	AC074386 Homo sapi
85	24.8	52.8	68130	3	AC024757 Zebrafish
86	24.8	52.8	85347	2	AC017560 Drosophila
87	24.8	52.8	89261	9	AC098863 Homo sapi
88	24.8	52.8	91268	8	AP004895 Lotus cor
89	24.8	52.8	99877	1	CP000003 Continuation (19 of
90	24.8	52.8	110000	2	AC151905 Mus muscu
91	24.8	52.8	122013	3	AC092399 Drosophila
92	24.8	52.8	122013	3	AC092399 Drosophila

```

c 93 24.8 52.8 128621 8 AC147002 Medicago
94 24.8 52.8 130030 9 AC004889 Homo sapi
95 24.8 52.8 132724 9 AC068713 Homo sapi
c 96 24.8 52.8 146059 5 AL929549 Zebrafish
97 24.8 52.8 151770 9 AC002066 Homo sapi
98 24.8 52.8 152467 2 AC140023 Medicago
99 24.8 52.8 152481 2 AC151210 Bos tauru
c 100 24.8 52.8 154904 9 AL162272 Human DNA

```

## ALIGNMENTS

```

RESULT 1
AR288266 LOCUS AR288266 47 bp DNA linear PAT 12-JUN-2003
DEFINITION Sequence 1 from patent US 6537751.
ACCESSION AR288266
VERSION AR288266.1 GI:31675550
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Cohen,D., Chumakov,I. and Blumenfeld,M.
TITLE Biallelic markers for use in constructing a high density
JOURNAL Patent: US 6537751-A 1 25-MAR-2003;
FEATURES
source
1..47
Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"

```

```

ORIGIN
Query Match 99.1%; Score 46.6; DB 6; Length 47;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CCACAGGCTTGATTGAAATAAATTTGATCCATTTTCAAAATTTT 47
Db 1 CCACAGGCTTGATTGAAATAAATTTGATCCATTTTCAAAATTTT 47

```

```

RESULT 2
AF257499/c LOCUS AF257499 93153 bp DNA linear HTG 26-JUL-2002
DEFINITION Homo sapiens chromosome 8 clone GSI-179c23 map 8p23.1, WORKING
DRAFT SEQUENCE, 7 unordered pieces.
AF257499
AF257499.3 GI:14327872
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_CANCELLED.
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 93153)
Lehmann,R., Menzel,U., Polley,A., Reichwald,K., Schudy,A.,
Sidiqul,R., Taudien,S., Wen,G., Schlegelberger,B., Siebert,R.,
Rosenthal,A. and Platzner,M.
Chromosome 8 genomic sequence
Unpublished
2 (bases 1 to 93153)
Schudy,A., Schillabel,M., Baumgart,C., Menzel,U., Schattevoy,R. and
Rosenthal,A.
Direct Submision
Submitted (18-Apr-2000) Genome Analysis, Institute of Molecular
Biotechnology, Butenbergrasse 11, Jena 07745, Germany
On Jun 7, 2001 this sequence version replaced gi:8151909.

```

```

TITLE JOURNAL
REFERENCE
AUTHORS
COMMENT
-----
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/

```

```

Contact: gscj-submit@genome.imb-jena.de
-----
Project Information
Center Project name: H127
Center Clone name: GSI-179c23
-----
Summary Statistics
Sequencing vector: M13; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 84498 bases at least Q40
Consensus quality: 89860 bases at least Q20
Quality coverage: 5,12 in Q20 bases; sum-of-contigs

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1 6411: contig of 6411 bp in length
* 6412 6511: gap of unknown length
* 6512 52674: contig of 46163 bp in length
* 52675 52774: gap of unknown length
* 52775 66702: contig of 13928 bp in length
* 66802 66802: gap of unknown length
* 66803 73537: contig of 6835 bp in length
* 73638 73737: gap of unknown length
* 73738 79228: contig of 5491 bp in length
* 79229 79328: gap of unknown length
* 82459 82459: contig of 3131 bp in length
* 82460 82559: gap of unknown length
* 82560 93153: contig of 10594 bp in length.

```

```

FEATURES
source
1..93153
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p23.1"
/clone="GSI-179c23"

```

## ORIGIN

```

Query Match 99.1%; Score 46.6; DB 2; Length 93153;
Best Local Similarity 97.9%; Pred. No. 5.5e-05;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CCACAGGCTTGATTGAAATAAATTTGATCCATTTTCAAAATTTT 47
Db 56629 CCACAGGCTTGATTGAAATAAATTTGATCCATTTTCAAAATTTT 56583

```

```

RESULT 3
AC011835/c LOCUS AC011835 182151 bp DNA linear PRI 11-DEC-2001
DEFINITION Homo sapiens chromosome 8, clone RP11-143D15, complete sequence.
AC011835
AC011835.8 GI:17488682
HTG.
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 182151)
Nusbaum,C. and Lander,E.
Birren,B., Linton,L.,
Unpublished
2 (bases 1 to 182151)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Collins,S., Collins,S., Collymore,A.,
Cooke,P., DeRellano,K., Dewar,K., Domingo,M., Donelan,L., Doyle,M.,

```

```

TITLE JOURNAL
REFERENCE
AUTHORS
COMMENT
-----
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/

```

Pereira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, T., Lehoczyk, J., Lieu, C., Locke, K., Macdonald, P., Margulis, N., McEwan, P., McQuirk, A., McKernan, K., McLaughlin, J., Meltrim, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Neill, D., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testfay, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE  
JOURNAL

Submitted (15-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 182151)

REFERENCE  
AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E. A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckgalter, B., Brown, A., Camarata, J., Campio, A., Chang, J., Chazaro, B., Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Laroque, K., Lamazares, R., Landers, T., Lehoczyk, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPherson, R., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Riese, C., Rogov, P., Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testfay, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zaitoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (04-SEP-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 182151)

REFERENCE  
AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E. A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckgalter, B., Brown, A., Camarata, J., Campio, A., Chang, J., Chazaro, B., Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Laroque, K., Lamazares, R., Landers, T., Lehoczyk, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPherson, R., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Riese, C., Rogov, P., Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testfay, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zaitoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL

Submitted (11-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Dec 11, 2001 this sequence version replaced gi:15426061.

## COMMENT

All repeats were identified using RepeatMasker:  
Smit, A. F. A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu

Contact: sequence submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: 12584  
Center clone name: 143\_D\_15

## FEATURES

source

Location/Qualifiers  
1. 182151  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-133D15"  
/clone\_lib="RP11-11 Human Male BAC"  
complement(512..597)  
/rpt\_family="Charliea"  
complement(610..913)  
/rpt\_family="Charliea"  
2178..2258  
/rpt\_family="MIR3"  
2276..2406  
/rpt\_family="MIR"  
4083..4241  
/rpt\_family="MIR3"  
5608..5891  
/rpt\_family="AluX"  
6464..6775  
/rpt\_family="AluY"  
6910..7023  
/rpt\_family="L2"  
8176..8197  
/rpt\_family="AT\_rich"  
9419..9485  
/rpt\_family="MIR"  
10035..10435  
/rpt\_family="MER47B"  
complement(11642..11929)  
/rpt\_family="AluSq"  
12215..12529  
/rpt\_family="AluX"  
12900..13193  
/rpt\_family="AluX"  
14043..14393  
/rpt\_family="(TA)n"  
14405..14503  
/rpt\_family="(TA)n"  
14532..14965  
/rpt\_family="(TA)n"  
complement(15222..15800)  
/rpt\_family="Tigger7"  
15675..15679  
/note="30 qual SNGC region"  
complement(15801..16106)  
/rpt\_family="AluY"  
15827..15836  
/note="30 qual SNGC region"  
complement(16107..16215)  
/rpt\_family="Tigger7"  
18288..18598  
/rpt\_family="AluX"  
20575..20598  
/rpt\_family="AT\_rich"  
20830..20858  
/rpt\_family="(CAT)n"  
complement(21041..21334)  
/rpt\_family="AluX"  
complement(22065..22181)  
/rpt\_family="LIM4"  
complement(22365..22545)  
/rpt\_family="LIMC1"  
complement(22630..22944)  
/rpt\_family="AluX"  
23695..24210

```

repeat_region /rpt_family="MERIA"
24726..24818 /rpt_family="MIR"
repeat_region complement(25222..25545)
/rpt_family="LIMBS"
repeat_region complement(25660..25726)
/rpt_family="MIR"
repeat_region 25905..26202 /rpt_family="AluSx"
repeat_region 27319..27425 /rpt_family="AluSx"
repeat_region 27426..27522 /rpt_family="(TTAA)n"
repeat_region 27559..28668 /rpt_family="AT_rich"
repeat_region 28750..28775 /rpt_family="L1PA5"
repeat_region /rpt_family="(A)n"
23067..29212 /rpt_family="AT_rich"
repeat_region 30630..31560 /rpt_family="L2"
repeat_region 31593..31621 /rpt_family="AT_rich"
repeat_region 33828..33887 /rpt_family="(GA)n"
repeat_region complement(34814..35105) /rpt_family="AluSg"
repeat_region complement(35165..36274) /rpt_family="L1ME1"
repeat_region complement(36275..36531) /rpt_family="AluSx"
repeat_region complement(36532..37050) /rpt_family="L1ME1"
repeat_region complement(37063..37228) /rpt_family="L1ME1"
repeat_region complement(37715..37801)

```

Query Match 99.1%; Score 46.6; DB 9; Length 182151;  
 Best Local Similarity 97.9%; Pred. No. 4.9e-05;  
 Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACGAGCTTGTAGAAATAAATTTGATCACCATTTCGAATTTT 47  
 Db 31661 CCACGAGCTTGTAGAAATAAATTTGATCACCATTTCGAATTTT 31615

RESULT 4  
 LOCUS AC016065/c 185463 bp DNA linear PRI 01-DEC-2001  
 DEFINITION Homo sapiens chromosome 8, clone RP11-115C21, complete sequence.  
 ACCESSION AC016065  
 VERSION AC016065.14 GI:17223348  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.  
 1 (bases 1 to 185463)  
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
 Homo sapiens chromosome 8, clone RP11-115C21  
 Unpublished  
 2 (bases 1 to 185463)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
 Baldwin,J., Barna,N., Beckert,R., Boguslavsky,L., Boukhgalter,B.,  
 Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,  
 Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donegan,L., Doyle,M.,  
 Ferrelia,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,  
 Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
 Howland,J.C., Johnson,R., Jones,C., Kam,L., Karatas,A., Klein,J.,  
 Lehoczy,J., Lieu,C., Locke,K., MacDonald,P., Margulis,N.,  
 McEwan,P., McGurk,A., McKernan,K., McLaughlin,D., Meltrin,J.,  
 Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
 Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,

Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talmas,J.,  
 Testfay,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
 Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (20-NOV-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 185463)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,  
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,  
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,  
 Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S.,  
 Ferrelia,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
 Jones,C., Kamat,A., Karatas,A., Kelle,C., LaRoque,K.,  
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,  
 Maclean,C., MacDonald,P., Major,J., Margulis,N., Matthews,C.,  
 McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrum,J.,  
 Menus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
 Raymond,C., Retts,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnpack,R.,  
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Struss,N., Subramanian,A., Talamas,J., Testfay,S., Theodore,J.,  
 Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zaitoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE  
 JOURNAL  
 COMMENT  
 Direct Submission  
 Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Dec 1, 2001 this sequence version replaced gi:16924135.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIRB  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L4199  
 Center clone name: 115\_C\_21  
 -----  
 FEATURES  
 source  
 1..185463  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /map="8"  
 /clone="RP11-115C21"  
 /clone\_11b="RPCT-11 Human Male BAC"  
 29..326  
 /rpt\_family="AluSx"  
 repeat\_region 1443..1549 /rpt\_family="(TTAA)n"  
 repeat\_region 1550..1646 /rpt\_family="AT\_rich"  
 repeat\_region 1683..2792 /rpt\_family="L1PA5"  
 /rpt\_family="L1PA5"  
 1722..1729  
 /note="c30 qual SNGL region"  
 2874..2898  
 /rpt\_family="(A)n"  
 repeat\_region 3190..3263 /rpt\_family="AT\_rich"  
 repeat\_region 3268..3333 /rpt\_family="(TTA)n"  
 repeat\_region 4743..5540 /rpt\_family="L2"  
 repeat\_region 5712..5740



```
repeat_region /rpt_family="AT_rich"
7947. .8004
/rpt_family="GA)n"
repeat_region complement(8931. .9222)
/rpt_family="AluSg"
repeat_region complement(9243. .10391)
/rpt_family="L1ME1"
repeat_region complement(10392. .10648)
/rpt_family="AluSx"
repeat_region complement(10649. .1167)
/rpt_family="L1ME1"
repeat_region complement(11180. .11345)
/rpt_family="L1ME1"
repeat_region complement(11832. .11937)
/rpt_family="Trigger7"
repeat_region complement(11931. .12426)
/rpt_family="Trigger7"
repeat_region 13849. .14153
/rpt_family="AluSg"
repeat_region 14625. .14878
/rpt_family="AluY"
repeat_region complement(14984. .15284)
/rpt_family="AluJo"
repeat_region 15303. .15386
/rpt_family="MIR3"
repeat_region 15495. .15540
/rpt_family="A-rich"
repeat_region 17196. .17230
/rpt_family="AT_rich"
repeat_region complement(17808. .18120)
/rpt_family="AluSg"
repeat_region complement(18195. .19089)
/rpt_family="MERS4B"
repeat_region 19123. .19232
/rpt_family="MER31-int"
repeat_region 19292. .19418
/rpt_family="MER31-int"
repeat_region complement(19421. .19549)
/rpt_family="FLAM_C"
repeat_region complement(19727. .19793)
/rpt_family="Alu"
repeat_region complement(19832. .19882)
/rpt_family="Char1el1"
repeat_region 20637. .20946
/rpt_family="AluY"
repeat_region 21160. .21499
/rpt_family="MER31-int"
repeat_region 21500. .21807
/rpt_family="AluSx"
repeat_region 21808. .21913
/rpt_family="MER31-int"
repeat_region 22546. .23228
/rpt_family="MER31-int"
repeat_region 23683. .23989
/rpt_family="AluSx"
repeat_region complement(25070. .25360)
/rpt_family="AluJo"
repeat_region 26370. .26675
/rpt_family="AluJo"
repeat_region 27392. .27528
/rpt_family="AluSx"
repeat_region 27525. .27660
/rpt_family="AluSx"
repeat_region 27661. .27963
/rpt_family="AluSc"
repeat_region 27964. .28294
/rpt_family="AluSx"
repeat_region 28295. .28350
/rpt_family="TGGA)n"
repeat_region complement(28719. .28842)
/rpt_family="MLT1L"
repeat_region complement(29249. .29551)
/rpt_family="AluSx"
```

```
repeat_region complement(31417. .31677)
/rpt_family="MIR"
repeat_region complement(31868. .32002)
/rpt_family="L2"
repeat_region 32628. .32941
/rpt_family="AluSx"
repeat_region 33756. .33887
/rpt_family="FLAM_C"
repeat_region 33905. .34179
/rpt_family="L1MED"
repeat_region 34838. .34991
/rpt_family="AluSg/x"
repeat_region 35321. .35630
/rpt_family="AluSg"
repeat_region complement(35769. .36076)
/rpt_family="AluY"
repeat_region 36214. .36385
/rpt_family="MERSB"
repeat_region complement(36646. .36956)
/rpt_family="AluSx"
repeat_region complement(40004. .40102)
/rpt_family="L2"
repeat_region 40547. .40576
/rpt_family="AT_rich"
repeat_region 41201. .41344
/rpt_family="L2"
repeat_region complement(41415. .41750)
/rpt_family="MLT1A2"
repeat_region 43260. .43526

Query Match 99.1%; Score 46.6; DB 9; Length 185463;
Best local similarity 97.9%; Pred. No. 4.9e-05;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACAGCTTGATTGAAATTAATTTGATCCATTTCGAATTTT 47
Db 5780 CCACAGCTTGATTGAAATTAATTTGATCCATTTCGAATTTT 5734

RESULT 5
AC018398 191377 bp DNA linear PRI 30-APR-2002
LOCUS Homo sapiens chromosome 8, clone RP11-16G12, complete sequence.
DEFINITION AC018398
ACCESSION AC018398
VERSION AC018398.10 GI:20198704
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 191377)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 8, clone RP11-16G12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 191377)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavsky,L., Bouckgalter,B., Brown,A., Castle,A., Colangelo,M.,
Collins,S., Collymore,A., Cooke,P., Dearlano,P., Dewar,K.,
Domino,M., Doyle,M., Fenesor,J., Ferreira,P., FitzHugh,M.,
Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B.,
Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L.,
Karacas,A., Klein,J., Landers,T., Lehoczy,U., Lieu,C., Locke,K.,
Macdonald,P., Margus,N., McEwan,P., McGurk,A., McKernan,K.,
Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., Peterson,K., Pierre,N., Pollara,V., Riley,R.,
Rothman,D., Roy,A., Santos,R., Severy,P., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirelli,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
```

REFERENCE  
AUTHORS

3 (bases 1 to 191377)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,  
Bouhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,  
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,  
Cook, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J. S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamit, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R.,  
Landers, T., Lehoczeky, J., Levine, R., Lindblad-Toh, K., Liu, G.,  
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,  
McCarthy, M., McEwan, P., McKernan, K., Melitrim, J., Menus, L.,  
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,  
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Olivier, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retter, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, U., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (19-APR-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 191377)

REFERENCE  
AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,  
Bouhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,  
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,  
Cook, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J. S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamit, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R.,  
Landers, T., Lehoczeky, J., Levine, R., Lindblad-Toh, K., Liu, G.,  
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,  
McCarthy, M., McEwan, P., McKernan, K., Melitrim, J., Menus, L.,  
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,  
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Olivier, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retter, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, U., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (30-APR-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 19, 2002 this sequence version replaced g1:20148178.  
All repeats were identified using RepeatMasker:  
Smt, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

## Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L3535  
Center clone name: 16\_G\_12

## FEATURES

## source

Location/Qualifiers  
1. 191377  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-16G12"

repeat\_region  
1. 221  
/rpt\_family="AluSg/x"  
complement(1277. 1428)  
/rpt\_family="L1ME1"  
repeat\_region  
1732. 1823  
/rpt\_family="CA)n"  
2184. 2238  
/note="PCR product sequence only"  
complement(2649. 2774)  
/rpt\_family="L2"  
repeat\_region  
2963. 3133  
/rpt\_family="TG)n"  
3167. 3208  
/rpt\_family="TG)n"  
3581. 3587  
/note="<30 qual SNGL region"  
complement(6732. 6840)  
/rpt\_family="L2"  
complement(8405. 8540)  
/rpt\_family="MIR3"  
8922. 9027  
/rpt\_family="L2"  
9061. 9174  
/rpt\_family="L2"  
9390. 9595  
/rpt\_family="L2"  
complement(9592. 9699)  
/rpt\_family="MIR"  
11091. 11333  
/rpt\_family="AluSx"  
11359. 11649  
/rpt\_family="AluSp"  
11765. 12100  
/rpt\_family="AluDb"  
12328. 12593  
/rpt\_family="AluDb"  
13535. 13813  
/rpt\_family="AluSp"  
13817. 13850  
/rpt\_family="CMA)n"  
13857. 14036  
/rpt\_family="MIR"  
17230. 17345  
/rpt\_family="L1ME1"  
17528. 17551  
/rpt\_family="AT\_rich"  
18471. 18757  
/rpt\_family="AluDo"  
21255. 21353  
/rpt\_family="L2"  
21420. 21484  
/rpt\_family="L1PA9"  
22294. 22456  
/rpt\_family="MIR"  
complement(22452. 22513)  
/rpt\_family="MIR"  
22668. 23024  
/rpt\_family="MT1A1"  
23390. 23520  
/rpt\_family="MERB5"  
23538. 23739  
/rpt\_family="MER2"  
23908. 24013  
/rpt\_family="MIR"  
24043. 24073  
/rpt\_family="TTTTG)n"  
complement(24079. 24356)  
/rpt\_family="AluSg"  
24877. 24952  
/rpt\_family="LTR67"  
27866. 28203  
/rpt\_family="AluSx"

```
repeat_region      28205..28272
                    /rpt_family="GA-rich"
repeat_region      28663..28682
                    /rpt_family="(7)n"
repeat_region      complement(28688..29140)
                    /rpt_family="MLT1C"
repeat_region      29551..29760
                    /rpt_family="L2"
repeat_region      complement(29782..30036)
                    /rpt_family="L1MED"
repeat_region      30200..30591
                    /rpt_family="MLT1B"
repeat_region      complement(30615..30910)
                    /rpt_family="L1MED"
repeat_region      31381..31634
                    /rpt_family="L2"
repeat_region      31679..31699
                    /rpt_family="AT-rich"
repeat_region      31939..31990
                    /rpt_family="L2"
repeat_region      33132..33161
                    /rpt_family="AT-rich"

Query Match      99.1%; Score 46.6; DB 9; Length 191377;
Best Local Similarity 97.9%; Pred. No. 4.9e-05;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CCACAGCGTTGATTAGAAATTAATTTGATCACCATTTCAAATTTT 47
Db      116244 CCACAGCGTTGATTAGAAATTAATTTGATCACCATTTCAAATTTT 116198

RESULT 6
LOCUS      AX087869      240825 bp      DNA      linear      PAT 17-MAR-2001
DEFINITION      Sequence 1 from Patent WO0114550.
ACCESSION      AX087869
VERSION      AX087869.1 GI:13396862
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1
AUTHORS      Barry,C., Chumakov,I. and Blumenfeld,M.
TITLE      Prostate cancer-related gene 3 (pg3) and biallelic markers thereof
JOURNAL      Patent: WO 0114550-A 1 01-MAR-2001;
GENSET      (FR)
FEATURES
source      1..240825
                    Location/Qualifiers
                    1..240825
                    /organism="Homo sapiens"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:9606"
                    1..2000
                    /note="5;regulatory region"
                    1823..1840
                    /note="5-390.pu"
                    1980..1998
                    /note="5-390-177.mis"
                    1987..2011
                    /bound_molecy="5-390-177.probe"
                    1999
                    /note="5-390-177 : polymorphic base G or C"
                    2000..2018
                    /note="5-390-177.mis complement"
                    2001..2079
                    /note="exon A"
                    2108..2125
                    /note="5-390.rp complement"
                    4559..4577
                    /note="5-391.pu"
                    4582..4600
                    /note="5-391-43.mis"

misc_binding      4589..4613
                    /bound_molecy="5-391-43.probe"
variation      4601
                    /note="5-391-43 : polymorphic base A or G"
                    4602..4620
                    /note="5-391-43.mis complement"
                    4627..4718
                    /note="exon B"
                    4891..4908
                    /note="5-391.rp complement"
                    10007..10025
                    /note="5-392.pu"
                    10115..10233
                    /note="exon C"
                    10209..10227
                    /note="5-392-222.mis"
                    10216..10240
                    /bound_molecy="5-392-222.probe"
                    10228
                    /note="5-392-222 : polymorphic base G or T"
                    10229..10247
                    /note="5-392-222.mis complement"
                    10267..10285
                    /note="5-392-280.mis"
                    10274..10298
                    /bound_molecy="5-392-280.probe"
                    10286
                    /note="5-392-280 : polymorphic base G or T"
                    10287..10305
                    /note="5-392-280.mis complement"
                    10370
                    /note="5-392-364 : insertion of G"
                    10411..10430
                    /note="5-392.rp complement"
                    26810..26897
                    /note="exon D"
                    31357..31471
                    /note="exon E"
                    34261..34404
                    /note="exon F"
                    37377..37466
                    /note="exon S"
                    39556..39574
                    /note="4-59.rp"
                    39704..40858
                    /note="exon T"
                    39877..39896
                    /note="4-58.rp"
                    39925..39943
                    /note="4-58-318.mis"
                    39932..39956
                    /bound_molecy="4-58-318.probe"
                    39944
                    /note="4-58-318 : polymorphic base G or T"
                    39945..39963
                    /note="4-58-318.mis complement"
                    39953..39970
                    /note="4-59.pu complement"
                    39954..39972
                    /note="4-58-289.mis"
                    39961..39985
                    /bound_molecy="4-58-289.probe"
                    39973
                    /note="4-58-289 : polymorphic base G or C"
                    39974..39992
                    /note="4-58-289.mis complement"
                    40242..40259
                    /note="4-58.pu complement"
                    41137..41154
                    /note="4-54.rp"
                    41366..41384
                    /note="4-54-199.mis"
                    41373..41397
```

```

primer_bind /bound_moiety="4-54-199.probe"
              41385..41403
              /note="4-54-180.mis"
variation 41385
primer_bind /note="4-54-199 : polymorphic base A or C"
              41386..41404
              /note="4-54-199.mis complement"
misc_binding 41392..41416
              /bound_moiety="4-54-180.probe"
variation 41404
primer_bind /note="4-54-180 : polymorphic base A or C"
              41405..41423
              /note="4-54-180.mis complement"
primer_bind 41564..41581
              /note="4-54.pu complement"
primer_bind 42122..42141
              /note="4-51.rp"
primer_bind 42213..42231
              /note="4-51-312.mis"
misc_binding 42220..42244
              /bound_moiety="4-51-312.probe"
variation 42232
primer_bind /note="4-51-312 : polymorphic base G or C"
              42233..42251
              /note="4-51-312.mis complement"
primer_bind 42526..42543
              /note="4-51.pu complement"
exon 50436..50545
              /note="exon G"
primer_bind 67289..67309
              /note="99-86.rp"
primer_bind 67456..67474
              /note="99-86-266.mis"
misc_binding 67463..67487
              /bound_moiety="99-86-266.probe"
variation 67475
primer_bind /note="99-86-266 : polymorphic base A or G"
              67476..67494
              /note="99-86-266.mis complement"
primer_bind 67724..67741
              /note="99-86.pu complement"
primer_bind 69182..69200
              /note="4-88.rp"
primer_bind 69502..69520
              /note="4-88-107.mis"
misc_binding 69509..69533
              /bound_moiety="4-88-107.probe"
variation 69521
primer_bind /note="4-88-107 : polymorphic base A or G"
              69522..69540
              /note="4-88-107.mis complement"
primer_bind 69609..69626
              /note="4-88.pu complement"
primer_bind 72696..72715
              /note="5-397.pu"
primer_bind 72819..72837
              /note="5-397-141.mis"
misc_binding 72826..72850
              /bound_moiety="5-397-141.probe"
variation 72838
primer_bind /note="5-397-141 : polymorphic base G or T"
              72839..72857
              /note="5-397-141.mis complement"
exon 72881..72918
              /note="exon H"
primer_bind 73099..73117
              /note="5-397.rp complement"
primer_bind 75858..75877
              /note="5-398.pu"
exon 75989..76151
              /note="exon I"
primer_bind 76041..76059
              /note="5-398-203.mis"

```

```

misc_binding 76048..76072
              /bound_moiety="5-398-203.probe"
variation 76060
primer_bind /note="5-398-203 : polymorphic base A or C"
              76061..76079
              /note="5-398-203.mis complement"
primer_bind 76289..76306
              /note="5-398.rp complement"
primer_bind 81006..81025
              /note="99-12738.pu"
primer_bind 81234..81252
              /note="99-12738-248.mis"
misc_binding 81241..81265
              /bound_moiety="99-12738-248.probe"
variation 81253

Query Match 99.1%; Score 46.6; DB 6; Length 240825;
Best Local Similarity 97.9%; Pred.No. 4.7e-05;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACAGGCTGATTGAATAAATTTGATCCACATTTCAATTTT 47
Db 83764 CCACAGGCTGATTGAATAAATTTGATCCACATTTCAATTTT 83810

RESULT 7
AX523960 240825 bp DNA linear PAT 21-NOV-2002
LOCUS AX523960
DEFINITION Sequence 1 from Patent WO02066641.
ACCESSION AX523960
VERSION AX523960.1 GI:25168889
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
AUTHORS Barry,C. and Chumakov,I.
TITLE Pg-3 and diallelic markers thereof
JOURNAL Patent: WO 02066641-A 1 29-AUG-2002;
GENSET (FR)

FEATURES
Location/Qualifiers
source 1..240825
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..2000
/note="5,regulatory region"
primer_bind 1823..1840
/note="5-390.pu"
primer_bind 1980..1998
/note="5-390-177.mis"
misc_binding 1987..2011
/note="5-390-177.probe"
variation 1999
/note="5-390-177 : polymorphic base G or C"
primer_bind 2000..2018
/note="5-390-177.mis complement"
exon 2001..2079
/note="exon A"
primer_bind 2108..2125
/note="5-390.rp complement"
primer_bind 4559..4577
/note="5-391.pu"
primer_bind 4582..4600
/note="5-391-43.mis"
misc_binding 4589..4613
/note="5-391-43.probe"
variation 4601
/note="5-391-43 : polymorphic base A or G"
primer_bind 4602..4620
/note="5-391-43.mis complement"
exon 4627..4718

```

```
/note="exon B"
primer_bind 4891..4908 /note="5-391.rp complement"
primer_bind 10007..10025
exon /note="5-392.pu"
primer_bind 10115..10233 /note="exon C"
misc_binding 10209..10227 /note="5-392-222.mis"
10216..10240 /bound_motif="5-392-222.probe"
variation 10228
primer_bind /note="5-392-222 : polymorphic base G or T"
10229..10247
primer_bind /note="5-392-222.mis complement"
10267..10285
misc_binding /note="5-392-280.mis"
10274..10298 /bound_motif="5-392-280.probe"
variation 10286
primer_bind /note="5-392-280 : polymorphic base G or T"
10287..10305
variation /note="5-392-280.mis complement"
10370
primer_bind /note="5-392-364 : insertion of G"
10411..10430
exon /note="5-392.rp complement"
26810..26897
exon /note="exon D"
31357..31471 /note="exon E"
34261..34404 /note="exon F"
37377..37466 /note="exon S"
39556..39574 /note="4-59.rp"
39704..40858 /note="exon T"
39877..39896 /note="4-58.rp"
39925..39943 /note="4-58-318.mis"
39932..39956 /bound_motif="4-58-318.probe"
39944
variation /note="4-58-318 : polymorphic base G or T"
39945..39963
primer_bind /note="4-58-318.mis complement"
39953..39970
primer_bind /note="4-59.pu complement"
39954..39972 /note="4-58-289.mis"
39961..39985 /bound_motif="4-58-289.probe"
39973
variation /note="4-58-289 : polymorphic base G or C"
39974..39992
primer_bind /note="4-58-289.mis complement"
40242..40259
primer_bind /note="4-58.pu complement"
41137..41154 /note="4-54.rp"
41366..41384 /note="4-54-199.mis"
41373..41397 /bound_motif="4-54-199.probe"
41385..41403 /note="4-54-180.mis"
41385
variation /note="4-54-199 : polymorphic base A or C"
41386..41404
primer_bind /note="4-54-199.mis complement"

misc_binding 41392..41416
variation /bound_motif="4-54-180.probe"
41404
primer_bind /note="4-54-180 : polymorphic base A or C"
41405..41423
primer_bind /note="4-54-180.mis complement"
41564..41581 /note="4-54.pu complement"
42122..42141 /note="4-51.rp"
42213..42231 /note="4-51-312.mis"
42220..42244 /bound_motif="4-51-312.probe"
42232
variation /note="4-51-312 : polymorphic base G or C"
42233..42251
primer_bind /note="4-51-312.mis complement"
42526..42543 /note="4-51.pu complement"
50436..50545 /note="exon G"
67289..67309 /note="99-86.rp"
67456..67474 /note="99-86-266.mis"
67463..67487 /bound_motif="99-86-266.probe"
67475
variation /note="99-86-266 : polymorphic base A or G"
67476..67494
primer_bind /note="99-86-266.mis complement"
67724..67741 /note="99-86.pu complement"
69182..69200 /note="4-88.rp"
69502..69520 /note="4-88-107.mis"
69509..69533 /bound_motif="4-88-107.probe"
69521
variation /note="4-88-107 : polymorphic base A or G"
69522..69540
primer_bind /note="4-88-107.mis complement"
69609..69626 /note="4-88.pu complement"
72698..72715 /note="5-397.pu"
72819..72837 /note="5-397-141.mis"
72826..72850 /bound_motif="5-397-141.probe"
72838
variation /note="5-397-141 : polymorphic base G or T"
72839..72857
exon /note="5-397-141.mis complement"
72881..72918 /note="exon H"
73099..73117 /note="5-397.rp complement"
75858..75877 /note="5-398.pu"
75989..76151 /note="exon I"
76041..76059 /note="5-398-203.mis"
76048..76072 /bound_motif="5-398-203.probe"
76060
variation /note="5-398-203 : polymorphic base A or C"
76061..76079
primer_bind /note="5-398-203.mis complement"
76289..76306
```

```

primer_bind      /note="5-398.rp complement"
                  81006..81025
primer_bind      /note="99-12738.pu"
                  81234..81252
misc_binding      /note="99-12738-248.mis"
                  81241..81265
                  /bound_molecly="99-12738-248.probe"
variation         81253

```

```

Query Match      99.1%; Score 46.6; DB 6; Length 240825;
Best Local Similarity 97.9%; Pred. No. 4.7e-05;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy              1 CCACAGGCTTGATTGAAATAAATTTGATCACCATTTCATAATTTT 47
Db              83764 CCACAGGCTTGATTGAAATAAATTTGATCACCATTTCATAATTTT 83810

```

# RESULT 8 HS591B8/c

```

LOCUS           142552 bp   DNA   linear   PRI 04-MAR-2003
DEFINITION      Human DNA sequence from clone RP4-591B8 on chromosome 1p13.1,
complete sequence.
ACCESSION       AL035410
VERSION         AL035410.7 GI:4775617
KEYWORDS        HTG.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens

```

```

REFERENCE       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE           1 (bases 1 to 142552)
JOURNAL         Whitehead, S.

```

```

COMMENT         Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
                  Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
                  humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
                  On May 11, 1999 this sequence version replaced gi:4678848.
                  ----- Genome Center
                  Center: Wellcome Trust Sanger Institute
                  Center code: SC
                  Web site: http://www.sanger.ac.uk
                  Contact: humquery@sanger.ac.uk

```

## REFERENCE

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone configs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1>

RP4-591B8 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

## FEATURES

```

Source          1..142552
                Location/Qualifiers
                1..142552
                /organism="Homo sapiens"

```

```

/mol_type="genomic DNA"
/db_xref="R2PD:RPCI704B08591"
/db_xref="taxon:9606"
/chromosome="1"
/map="p13.1"
/clone="RP4-591B8"
/clone_11b="RPCI-4"

```

```

Query Match      62.6%; Score 29.4; DB 9; Length 142552;
Best Local Similarity 80.5%; Pred. No. 14;
Matches 33; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy              6 GGCTTGATTGAAATAAATTTGATCACCATTTCATAATTTT 46
Db              87755 GGCTTGATTGAAATAAATTTGATCACCATTTCATAATTTT 87715

```

# RESULT 9 AC025631/c

```

LOCUS           160530 bp   DNA   linear   HTG 10-JUN-2000
DEFINITION      Home sapiens clone RP11-2104, WORKING DRAFT SEQUENCE, 4 unordered
pieces.
ACCESSION       AC025631
VERSION         AC025631.3 GI:8439937
KEYWORDS        HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens

```

```

REFERENCE       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE           1 (bases 1 to 160530)
JOURNAL         Birren, B., Linton, L., Nussbaum, C. and Lander, E.

```

```

COMMENT         Unpublished
                  2 (bases 1 to 160530)
                  Birren, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N.,
                  Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
                  Boguslavsky, L., Bouhgalter, B., Brown, A., Burkett, G.,
                  Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
                  Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
                  Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
                  Galagan, J., Gardyna, S., Glade, S., Goyette, M., Graham, L.,
                  Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
                  Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karakas, A.,
                  Klein, J., Labocque, K., Lamazares, R., Landers, T., Lehoczy, J.,
                  Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
                  McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McPheters, R.,
                  Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
                  Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
                  O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
                  Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
                  Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
                  Stange-Thomann, N., Stojanovic, N., Sudramanian, A., Talamas, J.,
                  Testa, S., Theodore, J., Tirrell, A., Travers, M., Trigglio, J.,
                  Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.,
                  Young, G., Zainoun, J., Zimmer, A. and Zody, M.

```

Direct Submission  
Submitted (12-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 10, 2000 this sequence version replaced gi:7960380.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

Project Information  
Center project name: I6305  
Center clone name: 21\_O\_4  
Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 158511 bases at least Q40  
Consensus quality: 159793 bases at least Q30  
Consensus quality: 160084 bases at least Q20  
Insert size: 158000; agarose-fp  
Insert size: 160230; sum-of-ctnigs  
Quality coverage: 6.9 in Q20 bases; agarose-fp  
Quality coverage: 6.8 in Q20 bases; sum-of-ctnigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1 4285: contig of 4285 bp in length  
\* 4386 4385: gap of 100 bp  
\* 4386 24998: contig of 20613 bp in length  
\* 24999 25098: gap of 100 bp  
\* 25099 74719: contig of 49621 bp in length  
\* 74720 74819: gap of 100 bp  
\* 74820 160530: contig of 85711 bp in length.  
Location/Qualifiers  
1.160530  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-2104"  
/clone\_1fb="RPCT-11 Human Male BAC"  
1.4285  
/note="assembly\_fragment"  
4386.24998  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left"  
25099.74719  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right"  
74820.160530  
/note="assembly\_fragment"

ORIGIN  
Query Match 62.6%; Score 29.4; DB 2; Length 160530;  
Best Local Similarity 80.5%; Pred. No. 14;  
Matches 33; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 6 GGCTTGATTAGAAATTAATTTGATCACCATTTCAAATT 46  
|||||  
Db 77337 GGCTTGCTTAGAAAAAAGTGAATTAACATTTTAAATT 77297

RESULT 10  
AL391060 266079 bp DNA linear HTG 10-JUL-2001  
LOCUS Homo sapiens chromosome 1 clone RP11-252M6, 3 unordered pieces.  
ACCESSION AL391060  
VERSION AL391060.14 GI:14348473  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_CANCELLED.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.  
1  
REFERENCE  
AUTHORS Thomas, D.  
TITLE Direct Submission  
JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone  
requests: clonerequest@sanger.ac.uk  
COMMENT On Jun 11, 2001 this sequence version replaced gi:13273825.

----- Genome Center  
Center: Sanger Centre  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquerry@sanger.ac.uk  
----- Project Information  
Center project name: BA252M6  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Sequencing vector: pLamId; L08752; 100% of reads  
Chemistry: Dye-terminator Big Dye; 83% of reads  
Chemistry: Dye-terminator ET-amersham; 16% of reads  
Consensus quality: 157888 bases at least Q40  
Consensus quality: 158496 bases at least Q30  
Consensus quality: 158748 bases at least Q20  
Insert size: 265879; sum-of-ctnigs  
Insert size: 156786; 6.7% error; agarose-fp  
Quality coverage: 3.77x in Q20 bases; sum-of-ctnigs  
Quality coverage: 6.40x in Q20 bases; agarose-fp

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1 159425: contig of 159425 bp in length  
\* 159426 159525: gap of 100 bp  
\* 159526 184277: contig of 24752 bp in length  
\* 184278 184377: gap of 100 bp  
\* 184378 266079: contig of 81702 bp in length.  
Location/Qualifiers  
1.266079  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/clone="RP11-252M6"  
/clone\_1fb="RPCT-11.1"  
1.159425  
/note="assembly\_fragment:02468"  
159526.184277  
/note="assembly\_fragment:02469"  
184378.266079  
/note="assembly\_fragment:02508"

ORIGIN  
Query Match 62.6%; Score 29.4; DB 2; Length 266079;  
Best Local Similarity 80.5%; Pred. No. 13;  
Matches 33; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 6 GGCTTGATTAGAAATTAATTTGATCACCATTTCAAATT 46  
|||||  
Db 30224 GGCTTGCTTAGAAAAAAGTGAATTAACATTTTAAATT 30264

RESULT 11  
AL356492 40479 bp DNA linear PRI 18-NOV-2000  
LOCUS Human DNA sequence from clone RP11-591I21 on chromosome 6, complete  
sequence.  
ACCESSION AL356492  
VERSION AL356492.10 GI:11229152  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.  
1 (bases 1 to 40479)  
REFERENCE  
AUTHORS Kay, M.  
TITLE Direct Submission



## JOURNAL

## COMMENT

Submitted (17-NOV-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
request: clonerequest@sanger.ac.uk  
On Nov 20, 2000 this sequence version replaced gi:11225749.  
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlap above, as we submit sequences with only a small overlap as described above.  
This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: EMBL, SW, SWISSPROT, Tr, TREMBL, Wp, WORMPEP, Information on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone configs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr6>

RP11-53121 is from the library RPCI-11.3 constructed at the Roswell Park Cancer Institute by the group of Peter de Jong. For further details see <http://bacpac.med.buffalo.edu/>

## VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-53121 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true left end of clone RP11-93K22 is at 40380 in this sequence. The true right end of clone RP11-801118 is at 100 in this sequence.

## Location/Qualifiers

## FEATURES

## source

```
1..40479
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-53121"
/clone_11b="RPCI-11.3"
1..1376
/note="L1PA13 repeat: matches 4799. .6155 of consensus"
repeat_region
1395..2170
/note="L1PA2 repeat: matches 1. .776 of consensus"
repeat_region
2166..7397
/note="L1PA2 repeat: matches 900. .6146 of consensus"
repeat_region
7398..10895
/note="L1MA4A repeat: matches 1471. .5001 of consensus"
repeat_region
10904..11118
/note="L1 repeat: matches 3684. .3909 of consensus"
repeat_region
11162..11393
/note="L1MA4A repeat: matches 5006. .5239 of consensus"
repeat_region
11386..12562
/note="L1MA4A repeat: matches 5122. .6300 of consensus"
repeat_region
14367..14708
/note="MLT1J repeat: matches 1. .351 of consensus"
repeat_region
15785..16013
/note="MIR repeat: matches 16. .247 of consensus"
repeat_region
16956..16999
/note="22 copies 2 mer tt 75% conserved"
repeat_region
18599..18729
/note="MIR repeat: matches 19. .153 of consensus"
repeat_region
19169..19766
/note="L2 repeat: matches 2060. .2672 of consensus"
repeat_region
21174..21354
/note="MER53 repeat: matches 4. .189 of consensus"
repeat_region
21765..21983
/note="AluSg/x repeat: matches 84. .305 of consensus"
repeat_region
21984..22286
/note="AluJo repeat: matches 14. .312 of consensus"
repeat_region
23259..23408
```

## ORIGIN

```
repeat_region /note="MIR repeat: matches 57. .214 of consensus"
24290..29248
repeat_region /note="L1PA5 repeat: matches 1177. .6141 of consensus"
29249..29535
repeat_region /note="MLT1J repeat: matches 104. .393 of consensus"
30752..31055
repeat_region /note="AluX repeat: matches 5. .295 of consensus"
31090..31115
repeat_region /note="13 copies 2 mer ac 92% conserved"
31166..31299
repeat_region /note="67 copies 2 mer tg 79% conserved"
32267..32575
repeat_region /note="AluY repeat: matches 1. .306 of consensus"
32643..32928
repeat_region /note="L1P repeat: matches 2952. .3240 of consensus"
32929..33366
repeat_region /note="L1PA8 repeat: matches 5727. .6162 of consensus"
33583..33885
repeat_region /note="AluX repeat: matches 1. .311 of consensus"
34284..35039
repeat_region /note="13 copies 2 mer ac 92% conserved"
35523..35817
repeat_region /note="L2 repeat: matches 896. .1674 of consensus"
35997..36035
repeat_region /note="AluX repeat: matches 1. .291 of consensus"
36036..36372
repeat_region /note="L2 repeat: matches 1701. .1737 of consensus"
36373..36938
repeat_region /note="MER7A repeat: matches 1. .346 of consensus"
37240..37509
repeat_region /note="L2 repeat: matches 1737. .2336 of consensus"
38616..39051
repeat_region /note="AluB repeat: matches 1. .293 of consensus"
/note="L2 repeat: matches 2248. .2750 of consensus"
```

```
Query Match 58.7%; Score 27.6; DB 9; Length 40479;
Best Local Similarity 75.0%; Pred. No. 67;
Matches 33; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 4 CAGGCTGATTGAAATTAAGTTGATCACCATTTCAAATTT 47
Db 13120 CAGCTGTGTAATAATGACTTATTTCCATTTCAAATTC 13163
```

```
RESULT 12
AC090635 70766 bp DNA linear HTG 04-MAR-2001
LOCUS Homo sapiens chromosome 8 clone RP11-297E18 map 8, LOW-PASS
DEFINITION
SEQUENCE SAMPLING.
```

```
AC090635
AC090635.1 GI:13194959
VERSION
HTG; HTGS PHASE0.
SOURCE
Homo sapiens (human)
```

```
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eulalia; Primates; Catarrhini; Homidae; Homo.
```

```
REFERENCE
1 (bases 1 to 70766)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 8, clone RP11-297E18
Unpublished
2 (bases 1 to 70766)
```

```
REFERENCE
AUTHORS
Bairren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Baatien,V., Boguslavsky,L., Boukhalter,B., Brown,A.,
Cammarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeRellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Fato,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardina,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heathord,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
Jehoczek,J., Levine,R., Liu,G., Maclean,C., MacDonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheters,R., Meldrum,J., Menus,L., Milnova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
```

TITLE  
JOURNAL  
COMMENT

O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,  
Rhinck, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roestli, M.,  
Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P.,  
Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,  
Travers, M., Trivis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,  
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J.,  
Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (04-MAR-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RW/RepeatMasker.html>

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

Project Information

Center project name: L10552  
Center clone name: 297\_E\_18

\*\*\*\*\* NOTE: This record contains 85 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1	732: contig of 732 bp in length
733	832: gap of 100 bp
833	1612: contig of 780 bp in length
1613	1712: gap of 100 bp
1713	2443: contig of 731 bp in length
2444	2543: gap of 100 bp
2544	3268: contig of 725 bp in length
3269	3368: gap of 100 bp
3369	4108: contig of 740 bp in length
4109	4208: gap of 100 bp
4209	4938: contig of 730 bp in length
4939	5038: gap of 100 bp
5039	5743: contig of 705 bp in length
5744	5843: gap of 100 bp
5844	6560: contig of 717 bp in length
6561	6660: gap of 100 bp
6661	7398: contig of 738 bp in length
7399	7498: gap of 100 bp
7499	8235: contig of 737 bp in length
8236	8335: gap of 100 bp
8336	9082: contig of 747 bp in length
9083	9182: gap of 100 bp
9083	9923: contig of 741 bp in length
9183	10023: gap of 100 bp
9924	10233: gap of 100 bp
10024	10750: contig of 727 bp in length
10751	10850: gap of 100 bp
10851	11583: contig of 733 bp in length
11584	11683: gap of 100 bp
11684	12415: contig of 732 bp in length
12416	12515: gap of 100 bp
12516	13237: contig of 722 bp in length
13237	13337: gap of 100 bp
13338	14070: contig of 733 bp in length
14071	14170: gap of 100 bp
14171	14892: contig of 722 bp in length
14892	14992: gap of 100 bp
14993	15713: contig of 721 bp in length
15714	15813: gap of 100 bp

15814	16541: contig of 728 bp in length
16542	16641: gap of 100 bp
16642	17389: contig of 748 bp in length
17390	17489: gap of 100 bp
17490	18235: contig of 746 bp in length
18236	18335: gap of 100 bp
18336	19092: contig of 757 bp in length
19093	19192: gap of 100 bp
19193	19936: contig of 744 bp in length
19937	20036: gap of 100 bp
20037	20782: contig of 746 bp in length
20783	20882: gap of 100 bp
20883	21624: contig of 742 bp in length
21625	21724: gap of 100 bp
21725	22454: contig of 730 bp in length
22455	22554: gap of 100 bp
22555	23284: contig of 730 bp in length
23285	23384: gap of 100 bp
23385	24117: contig of 733 bp in length
24118	24217: gap of 100 bp
24218	24937: contig of 720 bp in length
24938	25037: gap of 100 bp
25038	25755: contig of 718 bp in length
25756	25855: gap of 100 bp
25856	26605: contig of 750 bp in length
26606	26705: gap of 100 bp
26706	27434: contig of 725 bp in length
27435	27534: gap of 100 bp
27535	28264: contig of 730 bp in length
28265	28364: gap of 100 bp
28365	29094: contig of 730 bp in length
29095	29194: gap of 100 bp
29195	29944: contig of 750 bp in length
29945	30044: gap of 100 bp
30045	30780: contig of 736 bp in length
30781	30880: gap of 100 bp
30881	31605: contig of 725 bp in length
31606	31705: gap of 100 bp
31706	32424: contig of 719 bp in length
32425	32524: gap of 100 bp
32525	33279: contig of 755 bp in length
33280	33379: gap of 100 bp
33380	34108: contig of 729 bp in length
34109	34208: gap of 100 bp
34209	34944: contig of 736 bp in length
34945	35044: gap of 100 bp
35045	35802: contig of 758 bp in length
35803	35902: gap of 100 bp
35903	36642: contig of 740 bp in length
36643	36742: gap of 100 bp
36743	37473: contig of 731 bp in length
37474	37573: gap of 100 bp
37574	38315: contig of 742 bp in length
38316	38415: gap of 100 bp
38416	39131: contig of 716 bp in length
39132	39231: gap of 100 bp
39233	39964: contig of 733 bp in length
39965	40064: gap of 100 bp
40065	40795: contig of 731 bp in length
40796	40895: gap of 100 bp
40896	41618: contig of 723 bp in length
41619	41718: gap of 100 bp
41719	42457: contig of 719 bp in length
42438	42537: gap of 100 bp
42538	43265: contig of 728 bp in length
43266	43365: gap of 100 bp
43366	44094: contig of 729 bp in length
44095	44194: gap of 100 bp
44195	44928: contig of 734 bp in length
44929	45028: gap of 100 bp
45029	45754: contig of 726 bp in length
45755	45854: gap of 100 bp
45855	46588: contig of 734 bp in length

*	46589	46688: gap of 100 bp
*	46689	47423: contig of 735 bp in length
*	47424	47523: gap of 100 bp
*	47524	48244: contig of 721 bp in length
*	48245	48344: gap of 100 bp
*	48345	49070: contig of 726 bp in length
*	49071	49170: gap of 100 bp
*	49171	49901: contig of 731 bp in length
*	49902	50001: gap of 100 bp
*	50002	50733: contig of 732 bp in length
*	50734	50833: gap of 100 bp
*	50834	51549: contig of 716 bp in length
*	51550	51649: gap of 100 bp
*	51650	52393: contig of 744 bp in length
*	52394	52493: gap of 100 bp
*	52494	53232: contig of 739 bp in length
*	53233	53332: gap of 100 bp
*	53333	54082: contig of 750 bp in length
*	54083	54182: gap of 100 bp
*	54183	54921: contig of 739 bp in length
*	54922	55021: gap of 100 bp
*	55022	55760: contig of 739 bp in length
*	55761	55860: gap of 100 bp
*	55861	56584: contig of 724 bp in length
*	56585	56684: gap of 100 bp
*	56685	57409: contig of 725 bp in length
*	57410	57509: gap of 100 bp

Query Match  
Best Local Similarity 58.7%; Score 27.6; DB 2; Length 70766;  
Matches 33; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Qy 4 CAGCGTGTAGTAATAAATTTGATCACCATTTCATTTTCAATTTT 47  
Db 52913 CAGCTGTGTGAAATAGACTTATTTCATTTTCAATTTCAATTC 52956

RESULT 13  
AC006193  
LOCUS  
DEFINITION Arabidopsis thaliana chromosome I BAC F13011 genomic sequence.  
ACCESSION AC006193  
VERSION AC006193.3 GI:4733953  
KEYWORDS HTG.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM  
REFERENCE AUTHORS  
1 (bases 1 to 118335)  
Federpspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,  
Gonzalez,A., Kremenevskaja,I., Kim,C., Lenz,C., Li,J., Liu,S.,  
Luros,S., Schwartz,J., Shim,P., Toriumi,M., Vyotskaja,V.,  
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.  
Unpublished  
2 (bases 1 to 118335)  
Federpspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,  
Gonzalez,A., Kremenevskaja,I., Kim,C., Lenz,C., Li,J., Liu,S.,  
Luros,S., Schwartz,J., Shim,P., Toriumi,M., Vyotskaja,V.,  
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.  
Direct Submission  
Submitted (09-DEC-1998) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA  
3 (bases 1 to 118335)  
Federpspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,  
Gonzalez,A., Kremenevskaja,I., Kim,C., Lenz,C., Li,J., Liu,S.,  
Luros,S., Schwartz,J., Shim,P., Toriumi,M., Vyotskaja,V.,  
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
4 (bases 1 to 118335)  
Federpspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,  
Gonzalez,A., Kremenevskaja,I., Kim,C., Lenz,C., Li,J., Liu,S.,  
Luros,S., Schwartz,J., Shim,P., Toriumi,M., Vyotskaja,V.,  
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.  
Direct Submission  
Submitted (11-JUN-1999) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA  
5 (bases 1 to 118335)  
Federpspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,  
Altafi,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A.  
and Davis,R.W.  
Direct Submission  
Submitted (28-MAY-2000) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA  
On May 4, 1999 this sequence version replaced gi:4678192.  
COMMENT  
FEATURES  
source  
1..118335  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="I"  
/map="g1147"  
/clone="F13011"  
/gene="F13011.1"  
1092..2311  
/gene="F13011.1"  
join(1092..1225,1474..2311)  
/gene="F13011.1"  
/note="Hypothetical Protein"  
/codon\_start=1  
/protein\_id="AAD38245.1"  
/db\_xref="GI:5042406"  
/translation="MGQLIDPRKCTNPTONEWFEESTKLSAFOYTAQOITVMKIQ  
RPRFKMDKMQSSPEFTMTTRIPDPHSTGVREDAVDVCKFKLTENPYCCSOS  
QOHOHQRAFLIWDLNFIVFMESELSKAODEIKELKALDYERKARRAEIMTKLAK  
DVEERRMARAEAMQNKRLFKELSEKSEMYMKADLEEBERQMRRLAEVLRERYQAK  
IMDARLFLEKSLSELEBANROGERERNMMKPKILERACSSPARRCCEPQKRGINP  
PPRVRAIRSEKSKGSKLCEOKVQLKILLQKTPRCPLPLSSPP"  
complement(3352..3963)  
/gene="F13011.2"  
complement(3352..3963)  
/gene="F13011.2"  
/note="Hypothetical Protein"  
/codon\_start=1  
/protein\_id="AAD38246.1"  
/db\_xref="GI:5042407"  
/translation="MGNCLFCGLDGEEDDLIKVKSQGVLEFYSPTAGFVSHGS  
GHALFSAVDLMMKPLADHLVPCGSYVLPFIVSDELTKEFVGSCHVSNSELSAIF  
PYRMSIDNHRVLRKYSYTVDFSRNSHIRTROEKTRRRRTSSKGAIMKVMILINTER  
LIOLISDEGRTNELIESVRAVAKBETSSITSSSNFLSVQT"  
8107..9734  
/gene="F13011.3"  
join(8107..8140,8222..8356,8440..8489,8562..8893,  
8959..9062,9144..9219,9304..9734)  
/gene="F13011.3"  
/note="very similar to alcohol dehydrogenase; similar to  
gi|1143388 claes III ADH, glutathione-dependent  
formaldehyde dehydrogenase. [Arabidopsis thaliana],  
location of est 94M15T7 (gi|2757072) and 94M15XP (3')  
(gi|3450565)"  
/codon\_start=1  
/protein\_id="AAD38247.1"  
/db\_xref="GI:5042408"  
/translation="METGKVIYTCKAAYWAGAGEPLVDEYKVDPPQRLVRIIRLIFT  
SICHTDLSAMKGENDORAYPRILIGHEAAGIVESGVEEEMAGDHVLPITGCGD

CHVCKRDGANLCERFRVDPMKVMTDGTREFTSKONKPIYHFLNTSTFSEXTVIDS  
ACVLKVPLEPKIETSLSCGVSTDCMLNKGWGLAMVNDIOPASTVAFEGAGVL  
AVABRGARGAKETIGIDINPDKEFOLGRAGISEFINPKESADIAVHERMEITGEVGE  
YSEFCAGSIEALREARFLSTNSGCVATVMGVHASPOLLPIHHEMFQGSITNSVFG  
FKPKTQLPFFITQGLNLDLDFISHOLPFHDINEMQOLHOGKALRCLLHLH"  
/complement(10308..11958)  
/gene="F13011.4"  
/complement(10308..10729,10802..10883,11042..11125,  
11208..11384,11461..11617,11723..11958)  
/gene="F13011.4"  
/note="Location of ests B6H3T7 (gi|2414136), 62B10T7  
(gi|91264), 22B15T7A (gi|1217083). 89% identical to  
membrane related protein CP5 [Arabidopsis thaliana]"  
/codon\_start=1  
/product="Putative membrane related protein"  
/protein\_id="AAD38248.1"  
/db\_xref="GI:5042409"  
/translation="MMGLSSLVFTFPHSLGFSDDLIVFAPLMAVAVGLVGVWVRPK  
MAYLSKLNSPKPLNLQPLTSILKTSNHHVPSLSKEKGTGVTDDFRHLMLKVE  
VADGGPCWIMDMRSTPRTSYAMRRDPENGRPOVSRVTFEDATPEWVRDPWDER  
BKMDMLXSTSLERCKDTGTMVYVWAKPPEFCSDREYIIGRINDAGRPYCITK  
GVQYSPVRQNKPRRYDLYSSWCLRAVESKRGDMITCEVLLFHEHDMGLPWEIAK  
LVRQGMGAVAKKIEPGLRAYORAKAGGLSPSALMAHINTKVAEEFMNRGSIAE  
VTGDKPTGKNIPEKILVVGGAIALACTLDKGLTKAVIFGVARRFARMGRM"  
13535..15717  
/gene="F13011.5"  
/join(13535..13647,13808..14022,14535..14734,14807..15315,  
15402..15717)  
/gene="F13011.5"  
/note="Identical to Tubulin Alpha-1 Chain [Arabidopsis  
thaliana] (gi|135391) Location of ests 139020T7  
(gi|2762968), 214C14T7 (gi|1159108)."  
/codon\_start=1  
/product="alpha tubulin"  
/protein\_id="AAD38249.1"  
/db\_xref="GI:5042410"  
/translation="MREIISIHGAGIOWNSCWLTYCLRHGIOPDGMPSDTVGA  
CHDAFPTPSESSCOHVPRAVFLDEPVIDEVRGTGRQLPHEPQLISGEDANN  
FARGHTVREIVDTLERKLADNCTLGVLVNAVGGSLSLRLSRVD  
FGRKSLGFTIYSPQVSTAVVEPNYSVLSHLEHTVVLNDEALYDLCRSLD  
IRPPTYSNLRSLISQTSILTSLSRPDGAINDITEFQNLVPPRIHMLSYAVAVI  
SSAKYHEQFSEITTSVSEPSNMAKADPHRKGCCMLMAGVGVKDVNTAVAA  
IKAKRTIOPWDCPTGFKGIVQPSVPGDLAVORAVCMISNTTAVAFESRID  
HKFDLMSKAFVHMYVGEEMGESEARBDLALKYEEVSGAEDDEBGEY  
"  
/complement(16208..17133)  
/gene="F13011.6"  
/complement(16208..16386,17040..17133)  
/gene="F13011.6"  
/note="Unknown protein; Location of ests 247E21T7  
(gi|241341)"  
/codon\_start=1  
/protein\_id="AAD38250.1"  
/db\_xref="GI:5042411"  
/translation="MAEPRKATAVYVKMDLFBDDDEFEEFEINEDMLEKEVKEYSQ  
QMEDWDDDDVNDPSRQLRKELENGTIDK"  
/complement(17641..19522)  
/gene="F13011.7"  
/complement(17641..18251,18334..18833,19281..19522)  
/gene="F13011.7"  
/note="31% identical to Glucan endo-1,3-beta-glucosidase  
precursor ((1-3)-beta-glucan endohydrolase) [Triticum  
aestivum] (gi|1706551)"  
/codon\_start=1  
/product="Similar to glucan endo-1,3-beta-glucosidase  
precursor"  
/protein\_id="AAD38251.1"  
/db\_xref="GI:5042412"  
/translation="MATHLKPKTVVQMLKNNINKVLFDADETTMGALAGGLEVM  
VAIPNDQLKVTMSYERAKDMVRKAVRYNFDGVTTFVAVAGREPLKSYGSLFILT  
PALANIONALNEAGLNSGVKATVPLNADVDSPASNPARGPDDIIGQMTQIVD  
FIGKNNAPITINIPYLSLYGNDPFLNAYAFPGAEIPNIGLDYNNVDANPDLVS  
SLKAVGHGDMPIIVGEVGMPTGDKRANAGSAVRFYNGLLPRLGTNKGPLAPTYIEV

YLFGLIDEDAKSIAPGPFERNHGIKFKEDQPKFPIDLSCGQSGKFLIGANVPYLPNK  
WCTPFEARDLTKLANIDYACTFSQCTALGSSCNTLDANASVAFNMFPVQKQ  
DSACAFQGLATITTTQINISGQCNFPIQIVASASSFSQSSSVLVLIVWFLLSGMP  
"  
20970..22826  
/gene="F13011.8"  
/join(20970..21383,21459..21893,22659..22826)  
/gene="F13011.8"  
/note="Hypothetical Protein"  
/codon\_start=1  
/protein\_id="AAD38252.1"  
/db\_xref="GI:5042413"  
/translation="MASLISFSLPCKAVRSISAPOTQITTEKLEDFGRKGIKF  
SSNNIPWELKVRNSSLKLSLSDHVLVSRKYWKDGEFEVLYTVDGDSRGV  
GVYINGEERPKGSSVIGSCDMSVOTSDALDLOIELCTAGVLDITYVLSYVS  
MATLVYKRNKRPYTLKRGINSYLRFKRSAGIGLGCSCCPNPILSPPELLSP  
SEAMKAESSGWFGESEGERPGIWAVEDSVITLLEKMSIYGAPPAERLKAIVNTPS  
KEFTIDQDDKDMVRDVCIPRIDVQVWEATLLRVHRSNVNARSCRSLRDMERSN  
GNRA"  
/complement(24442..25986)  
QY 4 CAGGCTGATTAAGATAAATTTGATCACCATTTCAATTTT 47  
Db 70308 CAGGCTTGATTAAGCAATAGATTGATTTGATTTAGAACTTTT 70351  
Query Match 58.7%; Score 27.6; DB 8; Length 118335;  
Best Local Similarity 75.0%; Pred. No. 56;  
Matches 33; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
gene  
RESULT 14  
AC108046/c 177483 bp DNA linear PRI 20-MAR-2002  
LOCUS Homo sapiens BAC clone Rpl1-308K2 from 4, complete sequence.  
DEFINITION AC108046 AC024020  
AC108046.3 GI:19551179  
VERSION HTG.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 177483)  
AUTHORS Sulston,J.B. and Waterston,R.  
TITLE Toward a complete human genome sequence  
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)  
PUBMED 9847074  
REFERENCE 2 (bases 1 to 177483)  
AUTHORS Levy,A., Kozlowicz,A., Creason,K. and Pearman,C.  
TITLE The sequence of Homo sapiens BAC clone Rpl1-308K2  
JOURNAL Unpublished (2001)  
REFERENCE 3 (bases 1 to 177483)  
AUTHORS Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (24-JAN-2002) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
4 (bases 1 to 177483)  
Waterston,R.  
Direct Submission  
Submitted (20-MAR-2002) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Mar 20, 2002 this sequence version replaced gi:18642919.  
----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wucl.edu/gsc  
Contact: saplen@watscn.wucl.edu  
----- Summary Statistics  
Center project name: H\_NH00808K02  
Drafting Center: WMR

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P. Y., Zhao, B., Frengen, E., Tatenno, M., Catanesse, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBAC3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-696F12; the clone sequenced to the right is RP11-264F11. Actual start of this clone is at base position 1 of RP11-308K2; actual end is at base position 177483 of RP11-308K2.

The sequence of AC024020 has been incorporated into AC108046.

Location/Qualifiers

FEATURES  
source

```
1. 177483
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-308K2"
/clone_lib="RPCI-11"
1. 153
/rpc_family="L1"
repeat_region
/rpc_family="L1"
154. 464
/rpc_family="Alu"
repeat_region
465. 566
/rpc_family="L1"
repeat_region
567. 873
/rpc_family="Alu"
repeat_region
874. 956
/rpc_family="L1"
repeat_region
968. 1037
/rpc_family="L1"
repeat_region
1121. 1729
/rpc_family="L1"
repeat_region
1863. 2148
/rpc_family="MALR"
repeat_region
3785. 4013
/rpc_family="CT-rich"
repeat_region
4019. 4143
/rpc_family="L1"
repeat_region
4157. 4451
/rpc_family="Alu"
repeat_region
6910. 8125
/rpc_family="L1"
repeat_region
8171. 8290
/rpc_family="MIR"
repeat_region
8416. 8489
/rpc_family="L1"
```

```
repeat_region 8490. 8802
/rpc_family="Alu"
repeat_region 8803. 8960
/rpc_family="L1"
repeat_region 8961. 9012
/rpc_family="(TA)n"
repeat_region 9013. 9112
/rpc_family="L1"
repeat_region 9367. 9475
/rpc_family="MIR"
repeat_region 10800. 10822
/rpc_family="(TG)n"
repeat_region 10823. 11110
/rpc_family="Alu"
repeat_region 12603. 12692
/rpc_family="MIR"
repeat_region 12815. 12913
/rpc_family="(TATTA)n"
repeat_region 14271. 14396
/rpc_family="L2"
repeat_region 15058. 15345
/rpc_family="Alu"
repeat_region 15992. 16057
/rpc_family="L2"
repeat_region 16412. 16853
/rpc_family="L1"
repeat_region 17001. 17037
/rpc_family="AT-rich"
repeat_region 20274. 20395
/rpc_family="AT-rich"
repeat_region 21690. 21805
/rpc_family="MIR"
repeat_region 22201. 22328
/rpc_family="AT-rich"
repeat_region 22228. 22467
/rpc_family="L1"
repeat_region 22681. 22976
/rpc_family="Alu"
repeat_region 23210. 23552
/rpc_family="L1"
repeat_region 24520. 24549
/rpc_family="AT-rich"
repeat_region 24648. 24766
/rpc_family="MER1_type"
repeat_region 24777. 24898
/rpc_family="Alu"
repeat_region 24907. 25000
/rpc_family="MIR"
repeat_region 28464. 28913
/rpc_family="L1"
repeat_region 28941. 29039
/rpc_family="MER103"
repeat_region 29831. 29897
/rpc_family="L2"
repeat_region 30233. 30371
/rpc_family="MALR"
repeat_region 30372. 30841
/rpc_family="L1"
repeat_region 30855. 31383
/rpc_family="L1"
repeat_region 31391. 31779
/rpc_family="L1"
repeat_region 31810. 32339
/rpc_family="MALR"
repeat_region 32826. 32860
/rpc_family="AT-rich"
repeat_region 35299. 35393
/rpc_family="MIR"
repeat_region 35468. 35991
/rpc_family="MER2_type"
repeat_region 36009. 36168
/rpc_family="MER2_type"
repeat_region 36121. 36174
```

```

repeat_region      /rpt_family="MER2_type"
                    36365..36399
repeat_region      /rpt_family="(CA)n"
                    37093..37119
repeat_region      /rpt_family="(TTTG)n"
                    38314..38591
repeat_region      /rpt_family="L1"
                    39520..39623
repeat_region      /rpt_family="MIR"
                    39992..40066
repeat_region      /rpt_family="MIR"
                    40471..40635
repeat_region      /rpt_family="MIR"

```

```

Query Match      57.9%; Score 27.2; DB 9; Length 177483;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 32; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

Oy      8 CTTGATTAGAAATTAATTAATTTGATCACCATTTCATAATTT 47
Db      136383 CTTGTTAGAAATTAATTAATTTACTTACCATTTCCTTTT 136344

```

```

RESULT 15
BX469934      225586 bp      DNA      linear      VRT 25-SEP-2003
LOCUS      BX469934
DEFINITION      Zebrafish DNA sequence from clone CH211-209P5, complete sequence.
ACCESSION      BX469934
VERSION      BX469934.5      GI:35760789
KEYWORDS      HTG.
SOURCE      Danio rerio (zebrafish)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
      Cypriniformes; Cyprinidae; Danio.
      1 (bases 1 to 225586)
      Wood, J.
      Direct Submission
      Submitted (25-SEP-2003) Wellcome Trust Sanger Institute, Hinxton,
      Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
      zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
      On Sep 25, 2003 this sequence version replaced gi:35209448.
      ----- Genome Center
      Center: Wellcome Trust Sanger Institute
      Center code: SC
      Web site: http://www.sanger.ac.uk
      Contact: zfish-help@sanger.ac.uk
      -----

```

## COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep). Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those

```

beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, Mashu). For further information see
http://www.sanger.ac.uk/Projects/D\_rerio/fishmask.shtml CH211-209P5
is from a CHOR1-211 BAC library
VECTOR: pTABAC2.1.
      Location/Qualifiers
          1..225586
             /organism="Danio rerio"
             /mol_type="genomic DNA"
             /db_xref="taxon:7955"
             /clone="CH211-209P5"
             /clone_1fb="CHOR1-211"

```

## ORIGIN

```

Query Match      57.9%; Score 27.2; DB 5; Length 225586;
Best Local Similarity 76.2%; Pred. No. 67;
Matches 32; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

```

```

Oy      5 AGGCTTATTAGAAATTAATTAATTTGATCACCATTTCATAATTT 46
Db      70317 AGGATTGACTAGAAATTAATTAATTTGATCACCATTTCATAAT 70276

```

```

RESULT 16
BX908795      237042 bp      DNA      linear      HTG 20-FEB-2004
LOCUS      BX908795
DEFINITION      Danio rerio clone DKEX-92J12, WORKING DRAFT SEQUENCE, 10 unordered
      pieces.
ACCESSION      BX908795
VERSION      BX908795.4      GI:42733234
KEYWORDS      HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE      Danio rerio (zebrafish)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
      Cypriniformes; Cyprinidae; Danio.
      1 (bases 1 to 237042)
      Burton, J.
      Direct Submission
      Submitted (19-FEB-2004) Wellcome Trust Sanger Institute, Hinxton,
      Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
      zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
      On Feb 20, 2004 this sequence version replaced gi:4106643.
      ----- Genome Center
      Center: Wellcome Trust Sanger Institute
      Center code: SC
      Web site: http://www.sanger.ac.uk
      Contact: zfish-help@sanger.ac.uk
      ----- Project Information
      Center project name: ZK92J12
      ----- Summary Statistics
      Assembly program: XGAP4; Version 4.5
      Chemistry: Dye-terminator; 100% of reads
      Consensus quality: 232035 bases at least Q40
      Consensus quality: 232682 bases at least Q30
      Consensus quality: 233528 bases at least Q20
      Insert size: 236142; sum-of-coverage
      Insert size: 234948; 2.6% error; agarose-fp
      Quality coverage: 11.47% in Q20 bases; sum-of-coverage
      Quality coverage: 11.6% in Q20 bases; agarose-fp
      -----

```

## COMMENT

\* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 57482: contig of 57482 bp in length

\* 57483 57582: gap of 100 bp

\* 57583 109940: contig of 52358 bp in length

\* 109941 110040: gap of 100 bp

\* 110041 114134: contig of 4094 bp in length  
\* 114135 114234: gap of 100 bp  
\* 114235 149477: contig of 35243 bp in length  
\* 149478 149577: gap of 100 bp  
\* 149578 164819: contig of 15242 bp in length  
\* 164820 164919: gap of 100 bp  
\* 164920 185164: contig of 20245 bp in length  
\* 185165 185264: gap of 100 bp  
\* 185265 202375: contig of 17111 bp in length  
\* 202376 202475: gap of 100 bp  
\* 202476 204766: contig of 2291 bp in length  
\* 204767 204867: gap of 100 bp  
\* 204867 224261: contig of 19395 bp in length  
\* 224262 224361: gap of 100 bp  
\* 224362 237042: contig of 12681 bp in length.  
Location/Qualifiers  
source 1. 237042  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone\_idb="DantioKey"  
1. 57482  
/note="assembly\_fragment:03190  
fragment\_chain:1"  
misc\_feature 57583. 109940  
/note="assembly\_fragment:02229  
fragment\_chain:1"  
misc\_feature 110041. 114134  
/note="assembly\_fragment:00001  
fragment\_chain:1"  
misc\_feature 114235. 149477  
/note="assembly\_fragment:01609  
fragment\_chain:1"  
misc\_feature 149578. 164819  
/note="assembly\_fragment:00308  
fragment\_chain:1"  
misc\_feature 164920. 185164  
/note="assembly\_fragment:01211  
fragment\_chain:2"  
misc\_feature 185265. 202375  
/note="assembly\_fragment:00585  
fragment\_chain:2"  
misc\_feature 202476. 204766  
/note="assembly\_fragment:04241  
fragment\_chain:2"  
misc\_feature 204867. 224261  
/note="assembly\_fragment:00880  
fragment\_chain:3"  
misc\_feature 224362. 237042  
/note="assembly\_fragment:00061  
fragment\_chain:3"  
ORIGIN  
Query Match 57.9%; Score 27.2; DB 2; Length 237042;  
Best Local Similarity 76.2%; Pred. No. 66;  
Matches 32; Conservative 1; Mismatches 9; Indels 0; Gaps 0;  
Oy 5 AGGCTGATTGAAATTAATTAATTCACCATTTCAATT 46  
Db 141139 AGGATTGACTAGAAATTAATTCACCATTTCAATT 141098  
RESULT 17  
AL592222 160796 bp DNA linear ROD 19-SEP-2003  
LOCUS AL592222  
DEFINITION Mouse DNA sequence from clone RP23-366M19 on chromosome 11,  
complete sequence.  
ACCESSION AL592222  
VERSION AL592222.11 GI:34850734  
KEYWORDS HTG.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Submitted (19-SEP-2003) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk  
On Sep 19, 2003 this sequence version replaced gi:27899567.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep -----  
Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
-----  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest, except on the rare  
occasion of the clone being a YAC.  
RP23-366M19 is from the RPCI-23 Mouse BAC library  
constructed by the group of Pieter de Jong.  
For further details see http://www.chori.org/bacpac/home.htm  
VECTOR: pBACe3.6  
Sequence from the Mouse Genome Sequencing Consortium whole genome  
shotgun may have been used to confirm this sequence. Sequence data  
from the whole genome shotgun alone has only been used where it has  
a phred quality of at least 30.  
Location/Qualifiers  
source 1. 160796  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="11"  
/clone="RP23-366M19"  
/clone\_idb="RPCI-23"  
ORIGIN  
Query Match 56.2%; Score 26.4; DB 10; Length 160796;  
Best Local Similarity 71.7%; Pred. No. 13e+02;  
Matches 33; Conservative 1; Mismatches 12; Indels 0; Gaps 0;  
Oy 1 CCACAGGCTTGATTGAAATTAATTAATTCACCATTTCAATT 46  
Db 94768 CCACAGGCTTGCTTAAGTGTCTGTAGAGCCATTTCAATT 94723  
RESULT 18  
BX321875 169994 bp DNA linear VRT 01-SEP-2004  
LOCUS BX321875  
DEFINITION zebrafish DNA sequence from clone CH211-254C8 in linkage group 16,  
complete sequence.  
ACCESSION BX321875  
VERSION BX321875.7 GI:50724671  
KEYWORDS HTG.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Osteichthyes;



REFERENCE 1 (bases 1 to 169994)  
AUTHORS Cypriniformes; Cyprinidae; Danio.  
TITLE Pelan.S.  
JOURNAL Direct Submission  
Submitted (01-SEP-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk  
On Jul 27, 2004 this sequence version replaced gi:12398497.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: zfish-help@sanger.ac.uk  
-----  
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:  
EM: EMBL; SW: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep)  
Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.  
Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'dir' were identified by Rick Waterman (Stephen Johnson lab, Mashu). For further information see [http://www.sanger.ac.uk/Projects/D\\_rexio/fishmask.shtml](http://www.sanger.ac.uk/Projects/D_rexio/fishmask.shtml)  
is from a CHORI-211 BAC library  
VECTOR: pTRABAC2.1.  
Location/Qualifiers  
source  
1..169994  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="CH211-254C8"  
/clone\_11b="CHORI-211"  
ORIGIN  
Query Match 56.2%; Score 26.4; DB 5; Length 169994;  
Best Local Similarity 75.0%; Pred. No. 1.3e+02;  
Matches 33; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
Qy 3 ACAGGCTTGATTAGAAATTAAGTTGATCACCATTTCAAATTT 46  
Db 17472 ACATTTTGCATTATTATTAATCTGTCATCATTTTCAAAAT 17515  
RESULT 19  
BX936440 197773 bp DNA linear HTG 17-FEB-2004  
LOCUS BX936440  
DEFINITION Danio rerio clone DKEY-1K24, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 9  
uncovered pieces.  
ACCESSION BX936440  
VERSION BX936440.4 GI:42592599  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 197773)  
AUTHORS Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
TITLE Cypriniformes; Cyprinidae; Danio.  
JOURNAL Direct Submission  
Submitted (15-FEB-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk  
On Feb 17, 2004 this sequence version replaced gi:42557958.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: zfish-help@sanger.ac.uk  
-----  
Project Information  
Center project name: ZK1K24  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Chemistry: Dye-terminator; 100% of reads  
Consensus quality: 195294 bases at least Q40  
Consensus quality: 195829 bases at least Q30  
Consensus quality: 196182 bases at least Q20  
Insert size: 196973; sum-of-contigs  
Insert size: 186171; 1.5% error; agarose-fp  
Quality coverage: 6.96x in Q20 bases; sum-of-contigs Quality coverage: 7.36x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces \* is not known and their order in this sequence record is \* arbitrary. Gaps between the contigs are represented as \* runs of N, but the exact sizes of the gaps are unknown. \* This record will be updated with the finished sequence \* as soon as it is available and the accession number will \* be preserved.  
1 7723: contig of 7723 bp in length  
\* 1 7724: gap of 100 bp  
\* 7823: gap of 100 bp  
\* 7824 43716: contig of 35893 bp in length  
\* 43717 43816: gap of 100 bp  
\* 43817 97564: contig of 53748 bp in length  
\* 97565 97664: gap of 100 bp  
\* 97665 127663: contig of 29999 bp in length  
\* 127664 127763: gap of 100 bp  
\* 127764 137218: contig of 9455 bp in length  
\* 137219 137318: gap of 100 bp  
\* 137319 157224: contig of 19906 bp in length  
\* 157225 157324: gap of 100 bp  
\* 157325 168221: contig of 10897 bp in length  
\* 168222 168321: gap of 100 bp  
\* 168322 174805: contig of 6484 bp in length  
\* 174806 174905: gap of 100 bp  
\* 174906 197773: contig of 22868 bp in length.  
Location/Qualifiers  
source  
1..197773  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="DKEY-1K24"  
/clone\_11b="DanioKey"  
1..7723  
/note="assembly fragment:00077  
fragment\_chain:1"  
7824..43716  
/note="assembly fragment:00859  
fragment\_chain:1"  
43817..97564  
/note="assembly fragment:01613  
fragment\_chain:1"  
97665..127663  
/note="assembly fragment:01223  
fragment\_chain:1"  
127764..137218  
/note="assembly fragment:00163

```

misc_feature      fragment_chain:1"
                  137319..157224
                  /note="assembly_fragment:00366
                  fragment_chain:1"
misc_feature      157325..168221
                  /note="assembly_fragment:00260
                  fragment_chain:1"
misc_feature      168322..174805
                  /note="assembly_fragment:00016
                  fragment_chain:1"
misc_feature      174906..197773
                  /note="assembly_fragment:00603.0"

ORIGIN
Query Match      56.2%; Score 26.4; DB 2; Length 197773;
Best Local Similarity 71.7%; Pred. No. 1.2e+02;
Matches 33; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

Oy      2 CACAGGCTTGATTAGAAATTAASTTGATCACCATTTCAAATTTT 47
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      170985 CTCAGGCTTCATTAATAAAATACATCATCTGCATTTCAAATTTT 171030

RESULT 20
AC023173      203451 bp      DNA      linear      ROD 20-MAR-2002
LOCUS      Mus musculus chromosome 6 clone RP23-11G22 strain C57BL6/J,
DEFINITION      complete sequence.
ACCESSION      AC023173
VERSION      AC023173.3 GI:19551124
KEYWORDS      HTG.
SOURCE      Mus musculus (house mouse)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 203451)
AUTHORS      Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
Hachiguchi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Latic,P.,
Lee,Li,S.-Q., Legaapi,R., Maduro,Q.L., Maduro,V.B.,
Magallanes,E.H., Masiello,C., Maskeri,B., Mestrian,S.D.,
McCloskey,J.C., McDowell,J., Paguirigan,C., Pearson,R.,
Portnoy,M.E., Prasad,A., Schueler,M.G., Stancirpop,S., Thomas,J.W.,
Thomas,P.J., Touchman,J.W., Tsurugan,C., Vogt,J.L., Walker,M.A.,
Weberby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
TITLE      NISC Comparative Sequencing Initiative
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 203451)
AUTHORS      Green,E.D.
JOURNAL      Direct Submission
TITLE      Submitted (09-FEB-2000) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 203451)
Green,E.D.
JOURNAL      Direct Submission
TITLE      Submitted (20-MAR-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
On Mar 20, 2002 this sequence version replaced gi:12658031.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_mus@nigri.nih.gov
----- Project Information
Center project name: ya
Center clone name: 011G22

```

This sequence was finished as follows unless otherwise noted:  
all regions were double-stranded, sequenced with an  
alternate chemistry, or covered by high quality data  
(1.e., phred quality >= 30); an attempt was made to resolve  
all sequencing problems, such as compressions and repeats;

```

FEATURES
source
1..203451
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="6"
/clone="RP23-11G22"
/clone_1ib="RP23 mouse BAC library 23"
45177..45242
/note="single clone coverage"
49880..49881
/note="single clone coverage"
/note="low quality, single stranded/single chemistry
region"
122920..122971
/note="unresolved tandem repeat; consensus does not
reflect repeat; region sized by PCR"
123299..123345
/note="single clone coverage"
124835..124848
/note="single clone coverage"
153484..153530
/note="single clone coverage"
171106..171130
/note="single clone coverage"
195018..203451
/note="single clone coverage"
/note="clone overlaps with Genbank Accession Number
AC024950 (nucleotides 1-8433) clone RP23-208N9 (center
project name yb)"

misc_feature      122920..122971
misc_feature      123299..123345
misc_feature      124835..124848
misc_feature      153484..153530
misc_feature      171106..171130
misc_feature      195018..203451

ORIGIN
Query Match      56.2%; Score 26.4; DB 10; Length 203451;
Best Local Similarity 78.9%; Pred. No. 1.2e+02;
Matches 30; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Oy      10 TGATTAGAAATTAASTTGATCACCATTTCAAATTTT 47
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      111115 TTATTTAAATTAAGTTGATTCACATTTTAAATTTT 111152

RESULT 21
AC122835      218449 bp      DNA      linear      ROD 08-NOV-2003
LOCUS      Mus musculus BAC clone RP23-187118 from 3, complete sequence.
DEFINITION      AC122835
ACCESSION      AC122835
VERSION      AC122835.3 GI:23592205
KEYWORDS      HTG.
SOURCE      Mus musculus (house mouse)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 218449)
AUTHORS      Wang,C., Bielicki,L., Spalding,L. and Mangiapanello,L.
JOURNAL      The sequence of Mus musculus BAC clone RP23-187118
TITLE      Unpublished (2001)
REFERENCE      2 (bases 1 to 218449)
AUTHORS      Wilson,R.
JOURNAL      Sequencing of Mus musculus
TITLE      Unpublished (2001)
REFERENCE      3 (bases 1 to 218449)
AUTHORS      McPherson,J.D. and Waterston,R.H.
JOURNAL      Direct Submission
TITLE      Submitted (25-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 218449)
REFERENCE

```

AUTHORS	McPherson, J. D. and Waterson, R.H.
TITLE	Direct Submission
JOURNAL	Submitted (01-SEP-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	5 (bases 1 to 218449)
AUTHORS	McPherson, J. D. and Waterson, R.H.
TITLE	Direct Submission
JOURNAL	Submitted (08-OCT-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	6 (bases 1 to 218449)
AUTHORS	Wilson, R.
TITLE	Direct Submission
JOURNAL	Submitted (08-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT	On Oct 8, 2002 this sequence version replaced g1:22475871.

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

**MAPPING INFORMATION:**  
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

**SOURCE INFORMATION:**  
The PRC1-23 BAC Library has been constructed by Kazutoyo Oseagawa and Minko Tateono in the laboratory of Plietier de Jong and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Plietier de Jong and coworkers at <http://www.chori.org>

**NEIGHBORING SEQUENCE INFORMATION:**  
This sequence is the entire insert of the clone.

FEATURES	Location/Qualifiers
source	1. .218449

```
repeat__region      4921, -5047  
                    /rpt_family="MR"  
repeat__region      6201, -8301  
                    /rpt_family="LI"  
repeat__region      12177, -12233  
                    /rpt_family="All"  
repeat__region      12842, -13147  
                    /rpt_family="LI"  
repeat__region      13169, -13726  
                    /rpt_family="LI"  
repeat__region      14002, -14188
```

```
repeat_region      /rpt_family="B2"
                    15410..15496
repeat_region      /rpt_family="L2"
                    16050..16250
repeat_region      /rpt_family="B4"
                    16316..16844
repeat_region      /rpt_family="L1"
                    17939..18039
repeat_region      /rpt_family="B4"
                    19256..19343
repeat_region      /rpt_family="Alu"
                    20833..21026
repeat_region      /rpt_family="B2"
                    23277..23353
repeat_region      /rpt_family="Alu"
                    24094..24183
repeat_region      /rpt_family="B4"
                    25171..25354
repeat_region      /rpt_family="B2"
                    26413..26495
unrepeat           /note="Sequence derived from one plasmid subclone."
```

repeat_region	27759..27864	/rpt_family="Alu"
repeat_region	30322..30434	/rpt_family="Alu"
repeat_region	30353..30470	/rpt_family="B4"
repeat_region	30476..30685	/rpt_family="B2"
repeat_region	31557..31670	/rpt_family="Alu"
repeat_region	32025..32091	/rpt_family="ERV1"
repeat_region	33011..33137	/rpt_family="L1"
repeat_region	33179..33420	/rpt_family="L1"
repeat_region	33480..33672	/rpt_family="L1"
repeat_region	33678..33761	/rpt_family="L1"
repeat_region	33948..34358	/rpt_family="L1"
repeat_region	35496..35763	/rpt_family="B4"
repeat_region	38602..38729	/rpt_family="Alu"
repeat_region	40748..41002	/rpt_family="B4"
repeat_region	42010..42076	/rpt_family="ID"
repeat_region	44975..45052	/rpt_family="ID"
repeat_region	46218..46395	/rpt_family="MIR"
repeat_region	47806..47918	/rpt_family="B4"
repeat_region	50026..50225	/rpt_family="B2"
repeat_region	53346..53722	/rpt_family="MaLR"
repeat_region	55092..55281	/rpt_family="B2"
repeat_region	55474..55750	/rpt_family="L1"
repeat_region	57416..57966	/rpt_family="MaLR"
repeat_region	59233..59323	/rpt_family="L1"
repeat_region	59877..59999	/rpt_family="B4"
repeat_region	60000..60047	/rpt_family="ERV1"

```

repeat_region 62085..62387
/rpt_family="LI"
repeat_region 62841..62930
/rpt_family="MALR"
repeat_region 62931..63326
/rpt_family="MALR"
repeat_region 63327..63549
/rpt_family="MALR"
repeat_region 65251..66094
/rpt_family="LI"
repeat_region 66475..66867
/rpt_family="MALR"
repeat_region 67533..67936
/rpt_family="MALR"
repeat_region 68446..68824
/rpt_family="LI"
repeat_region 68925..69029
/rpt_family="LI"
repeat_region 69097..70572
/rpt_family="LI"
repeat_region 70566..70620
/rpt_family="LI"
repeat_region 70833..70932
/rpt_family="Alu"
repeat_region 71286..71402
/rpt_family="Alu"
repeat_region 71468..71864
/rpt_family="MALR"
repeat_region 71994..72035
/rpt_family="MALR"
repeat_region 74029..74162

```

Query Match 56.2%; Score 26.4; DB 10; Length 218449;  
 Best Local Similarity 71.7%; Pred No 1.2e+02;  
 Matches 33; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

2 CACAGCGCTTGATTAGAAATAAATTTGATCACCATTTCGAATTTT 47  
 Db 105833 CACAGCGCGATTGCAATTAACATTTCAAAATGCTTCAATTTT 105788

RESULT 22  
 AC119515/c 213417 bp DNA linear HTG 15-NOV-2002  
 LOCUS AC119515.5  
 DEFINITION Rattus norvegicus clone CH230-40602, \*\*\* SEQUENCING IN PROGRESS \*\*\*

AC119515  
 AC119515.5 GI:25012225  
 HTG: HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_ENRICHED.  
 Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE  
 AUTHORS 1 (bases 1 to 213417)  
 Munny,D,Marie., Metker,M,Dee., Abrazon,S., Adams,C., Alder,J.,  
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angiano,D.,  
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,  
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,  
 Biewald,K., Blais,J., Blankenburg,K., Blyth,P., Brown,M.,  
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,  
 Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,  
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,  
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,  
 Davila,M.L., Davis,C., Day-Carroll,L., De Anda,C., Dederich,D.,  
 Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,A., Divya,K.,  
 Diaper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,  
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,  
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,  
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,  
 Gevorgian,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,  
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,  
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,

TITLE  
 JOURNAL  
 AUTHORS  
 JOURNAL

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

COMMENT

Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,  
 Hollins,B., Howells,S., Hulik,S., Hune,J., Idlibird,D., Jackson,A.,  
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,  
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,  
 Kowals,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,  
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,  
 Lorenschewski,L., Louisedge,H., Lozada,R.U., Lu,X., Ma,D.,  
 Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,  
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,  
 Manning,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,  
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,  
 Morgan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Nair,L.,  
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,  
 Nwackelmech,O., Okwono,G., Olampunagson,A., Pal,S., Parks,K.,  
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,  
 Popper,F., Poindexter,A., Popovic,D., Prims,E., Pu,L.,  
 Puato,M., Qutroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,  
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,  
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.U.,  
 Sanders,M., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,  
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,  
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,  
 Steimle,M., Strong,R., Sutton,A., Sytek,A., Tabor,P., Taylor,C.,  
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Uemami,K.,  
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,  
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,  
 Williams,G., Willson,R., Wleczkyk,R., Wooden,H., Wolley,K.,  
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,  
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von  
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,  
 Weinstein,G., and Gibbs,R.A.  
 Direct Submission  
 2 (bases 1 to 213417)  
 Unpublished  
 2 (bases 1 to 213417)  
 Morley,K.C.  
 Submitted (28-APR-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 213417)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (15-NOV-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Nov 15, 2002 this sequence version replaced gi:23616941.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GQW  
 Center clone name: CH230-40602  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 199838 bases at least Q40  
 Consensus quality: 201193 bases at least Q20  
 Estimated insert size: 207854; sum-of-contigs estimation  
 Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

```

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
  1 213417: contig of 213417 bp in length.
    location/Qualifiers
      1.213417
      /organism="Rattus norvegicus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10116"
      /clone="CH230-40602"
      1.1798
      /note="wgs_end_extension"
      clone_end="5p6"
      6829..7711
      /note="clone boundary"
      clone_end="5p6"
      site_end="5p6"
      end_sequence="B2148299"
      206919..208667
      /note="wgs_contig"
      211434..213417
      /note="wgs_contig"

ORIGIN
Query Match      55.7%; Score 26.2; DB 2; Length 213417;
Best Local Similarity 79.5%; Pred. No. 1.4e+02;
Matches 31; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY      9 TTGATTAGAAATTAATTTGATCACCATTTCATTAATTTT 47
Db      147592 TTGTTAGAAATTAATTTGCAACATCATTTTGACATTTT 147554

RESULT 23
AC098034/c      217538 bp      DNA      linear      HTG 10-MAY-2003
LOCUS      AC098034
DEFINITION      Rattus norvegicus clone CH230-194A3, *** SEQUENCING IN PROGRESS
ACCESSION      AC098034
VERSION      AC098034.5 GI:30521938
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus
ORGANISM      Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
  1 (bases 1 to 217538)
  Muzny,D.,Matie., Metzker,M.,Lee., Adamzon,S., Adams,C., Alder,J.,
  Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D.,
  Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
  Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
  Bivwalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
  Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
  Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
  Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
  Cleveland,K., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
  Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
  Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
  Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
  Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
  Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
  Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
  Gebregiorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
  Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,

```

```

TITLE      JOURNAL
REFERENCE   JOURNAL
AUTHORS     HARVEY,Y., HAVLAK,P., HAWES,A., HENDERSON,N., HERNANDEZ,J.,
            HERNANDEZ,R., HINES,S., HLADUN,S.L., HODGSON,A., HOGUES,M.,
            HOLLINS,B., HOWELLS,S., HULYK,S., HUNE,J., IDLEBIRD,D., JACKSON,A.,
            JACKSON,L., JACOB,L., JIANG,H., JOHNSON,B., JOHNSON,R., JOLIVET,A.,
            KARPATHY,S., KELLY,S., KELLY,S., KHAN,Z., KING,L., KOVAR,C.,
            KOWIS,C., KRAFT,C.L., LABOW,H., LEVAN,J., LEWIS,L., LI,Z., LIU,J.,
            LIU,Y., LIU,Y., LONDON,P., LONGACRE,S., LOPEZ,J.,
            LORANUBHAW,L., LOUISEGE,H., LOZDO,R.J., LU,X., MA,J.,
            MAHESHWARI,M., MAHINDARNE,M., MAHMOUD,M., MALLOY,K., MANGUM,A.,
            MANGUM,B., MAPUA,P., MARTIN,K., MARTIN,R., MARTINEZ,E.,
            MAWHNEY,S., MCLOD,M.P., MCNEILL,T.Z., MEENEN,B.,
            MILOSAVLJEVIC,A., MINER,G., MINJA,E., MONTENAYOR,J., MOORE,S.,
            MORGAN,M., MORRIS,K., MORRIS,S., MUNDAAS,M., MURPHY,M., NAIR,L.,
            NANKERVIS,C., NEAL,D., NEWTON,N., NGUYEN,N., NORRIS,S.,
            NWAKEMENH,O., OKWONNU,G., OLARUNGBAOGON,A., PAL,S., PARKS,K.,
            PASTERNAK,S., PAUL,H., PEREZ,A., PEREZ,L., FRANKOCH,C.,
            PLOPPER,F., POINDEXTER,A., POPOVIC,D., PRIMUS,E., PU,L.-L.,
            PUZO,M., QUIROZ,J., RACHLIN,E., REEVES,K., REGIER,M.A., REIGH,R.,
            REILLY,B., REILLY,M., REN,Y., REUTER,M., RICHARDS,S., RIGGS,F.,
            RIVES,C., RODKEY,T., ROJAS,A., ROSE,M., ROSE,R., RUIZ,S.J.,
            SANDERS,M., SAVERY,G., SCHERER,S., SCOTT,G., SHATMAN,S., SHEN,H.,
            SHELTY,J., SHVARTSBEYN,A., SIBSON,I., SILTER,C.D., SMAILS,D.,
            SNEED,A., SODERGREN,E., SONG,X.-Z., SORELLE,R., SOGA,J.,
            STEINLE,M., STRONG,R., SUTTON,A., SVATEK,A., TABOR,P., TAYLOR,C.,
            TAYLOR,T., THOMAS,N., THOMAS,S., TINGEY,A., TRELOS,Z., USMANI,K.,
            VALAS,R., VERA,V., VILLASENA,D., WALDRON,L., WALKER,B., WANG,J.,
            WANG,Q., WANG,S., WARREN,J., WARREN,R., WEI,X., WHITE,F.,
            WILLIAMS,G., WILLSON,R., WLECZYK,R., WOODEN,H., WORLEY,K.,
            WRIGHT,D., WRIGHT,R., WU,J., YAKUB,S., YEN,D., YOON,L., YOON,V.,
            YU,F., ZHANG,J., ZHOU,X., ZHOU,X., ZHOU,S., DUNN,D., VON
            WEINSTECK,G. and Gibbs,R.A.
            Direct Submission
            Unpublished
            2 (bases 1 to 217538)
            Worley,K.C.
            Direct Submission
            Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            3 (bases 1 to 217538)
            Rat Genome Sequencing Consortium.
            Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            On May 10, 2003 this sequence version replaced gi:23096267.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.

            ----- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            ----- Project Information
            Center project name: GGAO
            Center clone name: CH230-194A3
            ----- Summary Statistics
            Assembly program: Atlas 3.0?
            Consensus quality: 192804 bases at least Q40
            Consensus quality: 195152 bases at least Q30
            Consensus quality: 197333 bases at least Q20
            Estimated insert size: 201938; sum-of-contigs estimation

```

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 4 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 65820: contig of 65820 bp in length  
 \* 65821 65920: gap of unknown length  
 \* 65921 213541: contig of 147621 bp in length  
 \* 213542 213641: gap of unknown length  
 \* 213642 215330: contig of 1889 bp in length  
 \* 215331 215630: gap of unknown length  
 \* 215631 217538: contig of 1908 bp in length.

Location/Qualifiers  
 1. 217538  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-194A3"  
 1. 1013  
 /note="wgs contig"  
 100203. 101969  
 /note="wgs\_contig"

ORIGIN  
 Query Match 55.7%; Score 26.2; DB 2; Length 217538;  
 Best Local Similarity 75.6%; Pred. No. 1.4e+02;  
 Matches 31; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

OY 7 GCTTGATTGAAATTAATTTGATCACCATTTCATTTT 47  
 ||||| :|||:|||||:|||||:|||||:|||||:|||||  
 Db 48092 GCATGTTGATCATCTTGTATCATCATTTTCAATCTT 48052

RESULT 24  
 AC107465 242655 bp DNA linear HTG 13-MAY-2003  
 LOCUS Rattus norvegicus clone CH230-86H10, \*\*\* SEQUENCING IN PROGRESS  
 AC107465  
 \*\*\* 3 unordered pieces.  
 AC107465 6 GI:30580831  
 HTG, HTGS\_PHASE1, HTGS\_DRAFT, HTGS\_ENRICHED.  
 Rattus norvegicus (Norway rat)  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 242655)  
 Muzny,D,Marie, Metzger,M, Lee, A, Abramson, S, Adams, C, Alder, J, Allen, C, Allen, H, Albrooks, S, Amin, A, Angilano, D, Anyalebechi, V, Aoyagi, A, Ayodeji, M, Baca, E, Baden, H, Baldwin, D, Bandaranaike, D, Barber, M, Barnstead, M, Benahmed, F, Bismail, K, Blair, J, Blankenburg, K, Blyth, P, Brown, M, Bryant, N, Buhay, C, Burch, P, Butrell, K, Calderon, E, Cardenas, V, Carter, K, Cavazos, I, Caesar, H, Center, A, Chacko, J, Chavez, D, Chen, G, Chen, R, Chen, Y, Chen, Z, Chu, J, Cleveland, C, Cockrell, R, Cox, C, Coyle, M, Cree, A, D'Souza, L, Davila, M, L, Davis, C, Davy-Carroll, L, De Anda, C, Dederich, D, Delgado, O, Denison, S, Deramo, C, Ding, Y, Dinh, H, Divya, K, Draper, H, Dugan-Rocha, S, Dunn, A, Durbin, K, Duval, B, Eaves, K, Egan, A, Escotto, M, Eugene, C, Evans, C, A, Falls, T, Fan, G, Fernandez, S, Finley, M, Flagg, N, Forbes, L, Foster, M, Foster, P, Fraser, C, M, G, Gant, R, Garcia, A, Garner, T, Garza, M, Geregorgis, E, Geer, K, Gill, R, Grady, M, Guerra, W, Guevara, W, Gunaratne, P, Haaland, W, Hamil, C, Hamilton, C, Hamilton, K, Harvey, Y, Havlak, P, Hawes, A, Henderson, N, Hernandez, J, Hernandez, R, Hines, S, Hladun, S, L, Hodgson, A, Hogues, M,

Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lotenshewa, L., Louissegh, H., Lozdo, R. J., Lu, X., Ma, J., Mangum, B., Mapua, P., Mahindaratne, M., Mahmoud, M., Mallory, K., Mangum, A., Manning, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Mlasev, J., A. A., Miner, G., Minja, E., Montanayor, J., Morgan, M., Morris, K., Morris, S., Mundaas, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwakoelamen, O., Okwou, G., Olarunpagaon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Poindexter, A., Popovic, D., Piums, E., Pu, L., Puazo, M., Quiroz, J., Rachin, E., Reeves, K., Regier, M. A., Reich, R., Rellis, B., Rellis, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. D., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Slason, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabors, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tinney, A., Trejos, P., Umani, K., Valas, R., Vera, V., Villasana, D., Walron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G., and Gibbs, R. A.

Direct Submission  
 Unpublished  
 2 (bases 1 to 242655)  
 Worley, K. C.

Direct Submission  
 Submitted (22-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 242655)  
 Rat Genome Sequencing Consortium.

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 On May 13, 2003 this sequence version replaced gi:22855692.  
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 Project Information  
 Center project name: G1XA  
 Center clone name: CH230-86H10  
 Summary Statistics  
 Assembly program: Atlas 3.0;  
 Consensus quality: 230159 bases at least Q40  
 Consensus quality: 232877 bases at least Q30  
 Consensus quality: 234876 bases at least Q20  
 Estimated insert size: 245378; sum-of-contigs estimation  
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 3 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 239139: contig of 239139 bp in length  
 \* 239140 239239: gap of unknown length  
 \* 239240 240581: contig of 1342 bp in length  
 \* 240582 240681: gap of unknown length  
 \* 240682 242655: contig of 1974 bp in length.

FEATURES  
 source  
 1. 242655  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-86H10"  
 1. 1966  
 /note="wgs\_contig"  
 misc\_feature  
 2365. 5223  
 /note="wgs\_contig"  
 237046. 239139  
 /note="wgs\_contig"

ORIGIN  
 Query Match 55.7%; Score 26.2; DB 2; Length 242655;  
 Best Local Similarity 79.5%; Pred. No. 1.4e+02;  
 Matches 31; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 9 TTGATTAGAAATTAATTTGATCACCATTTCATTTT 47  
 |||||  
 Db 232179 TTGTTAGAAATTAATTTGCAATCATTTT 23217

RESULT 25  
 AC131879/c  
 LOCUS  
 DEFINITION  
 AC131879 313096 bp DNA linear HTG 15-NOV-2002  
 Rattus norvegicus clone CH230-86H14, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\* 4 unordered pieces.  
 AC131879.5 GI:25007070  
 HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.  
 Rattus norvegicus (Norway rat)  
 ORGANISM  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 313096)  
 Muzny, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,  
 Allen, C., Allen, H., Alpbrosky, S., Amin, A., Anguiano, D.,  
 Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,  
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, W., Benedmed, F.,  
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
 Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,  
 Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Y., Chu, J.,  
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
 Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,  
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
 Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
 Gebregiorgis, B., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,  
 Guneratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
 Harrey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,  
 Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A.,  
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

Karpthy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
 Kowis, C., Kraft, C. L., Labow, H., Lavan, J., Lewis, L., Li, Z., Liu, J.,  
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
 Lounsbury, L., Lounsbury, H., Lozano, R. J., Lu, X., Ma, J.,  
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Mallory, K., Mangum, A.,  
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, B.,  
 Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,  
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
 Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L.,  
 Nankervis, C., Neal, D., Newcom, N., Nguyen, N., Norris, S.,  
 Nwackeleme, O., Okwodu, G., Olarinmaga, A., Pal, S., Parks, K.,  
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
 Plummer, F., Poldinger, A., Popovic, D., Prims, E., Pu, L.,  
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Register, M. A., Reid, R.,  
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,  
 Sanders, M., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,  
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D.,  
 Sneed, A., Sodergren, E., Song, X.-Z., Sorrell, R., Soza, J.,  
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemari, K.,  
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,  
 Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
 Williams, G., Willson, R., Wleczek, R., Wood, H., Worley, K.,  
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
 Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,  
 Weinstock, G. and Gibbs, R. A.

Unpublished  
 Direct Submission  
 2 (bases 1 to 313096)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (27-AUG-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 313096)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (15-NOV-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Nov 15, 2002 this sequence version replaced gi:23689082.

COMMENT  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (<http://www.hgsc.bcm.tmc.edu/projects/atlas/>). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

----- Genome Center -----  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information -----  
 Center project name: GSNZ  
 Center clone name: CH230-86H14  
 ----- Summary Statistics -----  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 22307 bases at least Q40  
 Consensus quality: 225821 bases at least Q30  
 Consensus quality: 228136 bases at least Q20  
 Estimated insert size: 232385; sum-of-contigs estimation  
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).



\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 252512: contig of 252512 bp in length  
\* 252513 252612: gap of unknown length  
\* 252613 266426: contig of 13814 bp in length  
\* 266427 266526: gap of unknown length  
\* 266527 310589: contig of 44033 bp in length  
\* 310590 310659: gap of unknown length  
\* 310660 313096: contig of 2437 bp in length.

FEATURES

source

1. .313096  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-86114"  
252613. .254069  
/note="wgs\_contig"

misc\_feature

ORIGIN

Query Match 55.7%; Score 26.2; DB 2; Length 313096;

Best Local Similarity 75.6%; Pred. No. 1.3e+02;

Matches 31; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 7 GCTTATTGAAATAAATTGATCAGCATTTTCAAAATTT 47

Db 47852 GCATGTTGATCATCACTTGTATCACCATTTCATATCTT 47812

Search completed: June 2, 2005, 10:20:34  
Job time : 1851 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2005, 08:12:23 / Search time 450 Seconds  
(without alignments)  
618.284 Million cell updates/sec

Title: US-10-349-143-1

Perfect score: 47

Sequence: 1 ccacagcgtcgtatagaat.....gatcacatttcaatttc 47

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: geneseqn1808:\*  
2: geneseqn1908:\*  
3: geneseqn2008:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46.6	99.1	47	3	AA265654
2	46.6	99.1	240823	10	AD69391
3	46.6	99.1	240825	4	AA24497
4	46.6	99.1	240825	6	AB081802
5	24.4	51.9	47108	6	ABK31511
6	24.2	51.5	993	9	ADA32589
7	24.2	51.5	4937	6	ABL07356
8	24.2	51.73	5173	6	ABL32951
9	23.6	50.2	27681	4	AA336498
10	23.6	50.2	27681	4	AA336497
11	23.6	50.2	27681	4	AAK85843
12	23.6	50.2	27681	10	AD647191
13	23.6	50.2	27681	10	AD647192
14	23.6	50.2	27681	13	AD108610
15	23.6	50.2	27681	13	AD108609
16	23.4	49.8	342	5	AAH33984
17	23.4	49.8	34917	4	AAK70686
18	23.2	49.4	564	10	ACF69221
19	23.2	49.4	939	6	ABN67694
20	23.2	49.4	1978	8	ACD13402

21	23.2	49.4	1978	13	ACN41054
22	23.2	49.4	2000	12	ADJ40809
23	23.2	49.4	2347	11	ADM02167
24	23.2	49.4	3212	4	AB128314
25	23.2	49.4	3468	4	AAK89062
26	23.2	49.4	40359	4	ABJ30382
27	23.2	49.4	61197	11	ACN45130
28	23.2	49.4	73742	12	ADQ97824
29	23.2	49.4	90104	4	AB112402
30	23.2	49.4	90435	12	ADQ59524
31	23.2	49.4	109661	12	ADQ97818
32	23.2	49.4	110000	10	ADF77343_09
33	23.2	49.4	110000	10	ACF67367_21
34	23.2	49.4	114693	8	AAD48308
35	23.2	49.4	118384	4	ABX56555
36	23.2	49.4	126001	12	ADH77123
37	23.2	49.4	243072	10	ACF65382
38	23.2	49.4	360	10	ACF70068
39	23.2	48.9	411	8	ABX54130
40	23.2	48.9	571	6	ABN60688
41	23.2	48.9	3874	4	AB111142
42	23.2	48.9	4254	4	AB111115
43	23.2	48.9	10329	6	ABJ34122
44	23.2	48.9	12291	4	AAK79265
45	23.2	48.9	12646	4	AB111114
46	23.2	48.9	50000	9	ADH16926
47	23.2	48.9	110000	10	ACF65385_2
48	23.2	48.9	110000	10	ACF67367_31
49	23.2	48.5	1506	2	AA767506
50	22.8	48.5	2751	2	AAK30434
51	22.8	48.5	2784	2	AA768270
52	22.8	48.5	28397	12	ADQ59407
53	22.8	48.5	31491	4	AD10203
54	22.8	48.5	31491	4	AAK63301
55	22.8	48.5	31491	6	AAK98863
56	22.6	48.1	360	5	AAK65972
57	22.6	48.1	459	2	AAV75449
58	22.6	48.1	642	8	ACF73906
59	22.6	48.1	701	3	AAK13788
60	22.6	48.1	726	4	AA194957
61	22.6	48.1	1209	2	AA783834
62	22.6	48.1	1209	2	AAV53355
63	22.6	48.1	2570	9	ACCS9915
64	22.6	48.1	10174	4	AB103298
65	22.6	48.1	24389	4	AB129006
66	22.6	48.1	28564	10	ADQ48909
67	22.6	48.1	73725	9	ADQ02690
68	22.6	48.1	73725	10	ADH72428
69	22.6	48.1	73725	10	ADH95938
70	22.6	48.1	110000	11	ACN44934_1
71	22.6	48.1	245531	13	ABD33022
72	22.6	48.1	322101	10	AAK58431
73	22.6	48.1	322101	12	AD135046
74	22.4	47.7	472	9	ACH37763
75	22.4	47.7	492	13	ADH92630
76	22.4	47.7	644	10	ADH17465
77	22.4	47.7	745	8	ACA04735
78	22.4	47.7	765	4	AAH08429
79	22.4	47.7	2012	12	ADQ20962
80	22.4	47.7	3634	13	ACN42674
81	22.4	47.7	3760	13	ACN42673
82	22.4	47.7	3761	13	ACN42672
83	22.4	47.7	3777	6	AB165195
84	22.4	47.7	3874	4	AB112274
85	22.4	47.7	3986	10	ADH76730
86	22.4	47.7	4003	12	ADQ24897
87	22.4	47.7	4013	10	ADH31412
88	22.4	47.7	4032	13	ACN42671
89	22.4	47.7	4688	5	ABV25850
90	22.4	47.7	4688	5	ABV25472
91	22.4	47.7	5085	8	ACF62791
92	22.4	47.7	5085	8	ACF62813
93	22.4	47.7	5087	13	ADH89577

```
c 94 22.4 47.7 5087 13 ADS89363 Ad889363 Oligonuc1
c 95 22.4 47.7 5087 13 ADS89303 Ad889303 Oligonuc1
c 96 22.4 47.7 5087 13 ADS89637 Ad889637 Oligonuc1
c 97 22.4 47.7 8085 4 AAs46479 Tumour su
c 98 22.4 47.7 8085 6 ABK33986 Human DNA
c 99 22.4 47.7 8085 8 ADA20374 Prostate
c 100 22.4 47.7 8085 8 ADA84181 Human ren
```

## ALIGNMENTS

```
RESULT 1
AAZ65654
ID AAZ65654 standard; DNA; 47 BP.
```

```
XX AAZ65654;
AC
XX 11-SEP-2001 (first entry)
DT
XX Human map-related biallelic marker SEQ ID NO.1.
DE
XX Human genome; biallelic marker; high density disequilibrium map;
KW genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW haplotyping; hybridisation; identification; characterisation; diagnosis;
KW single nucleotide polymorphism; SNP; ds.
```

```
XX Homo sapiens.
```

```
XX Key Location/Qualifiers
FH variation replace(24, C)
FT /*tag= a
```

```
FT /standard_name= "single nucleotide polymorphism"
```

```
XX WO954500-A2.
```

```
XX 28-OCT-1999.
```

```
XX 21-APR-1999; 99WO-1B000822.
```

```
XX 21-APR-1998; 98US-0082614P.
```

```
XX 23-NOV-1998; 98US-0109732P.
```

```
XX (GENE) GENSET.
```

```
XX Cohen D, Blumenfeld M, Chumakov I;
```

```
XX WPI; 2000-013267/01.
```

```
XX Novel biallelic markers used to construct a high density disequilibrium
XX map of the human genome.
```

```
XX Claim 1; Page 230; 2745PP; English.
```

```
XX AAZ65654 to AAZ69578 represent human biallelic markers from the present
XX invention, which contain a polymorphic base at position 24 of their
XX nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
XX primers for the biallelic markers. The biallelic markers of the invention
XX have a variety of uses: they can be used for high density mapping of the
XX human genome, and in complex association studies and haplotyping studies
XX which are useful in determining the genetic basis for disease states.
XX Compositions and methods of the invention can also be useful for the
XX identification of the targets for the development of pharmaceutical
XX agents and diagnostic methods, as well as the characterisation of the
XX differential efficacious responses to and side effects from
XX pharmaceutical agents acting on a disease as well as other treatment.
XX N.B. The SEQ ID NOS 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297 and
XX 3367, are not actually given a sequence in the Sequence Listing from the
XX present invention
```

```
XX Sequence 47 BP; 16 A; 8 C; 6 G; 17 T; 0 U; 0 Other;
```

```
XX Query Match 99.1%; Score 46.6; DB 3; Length 47;
```

```
Best Local Similarity 97.9%; Pred. No. 2.2e-06;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CCACGGCTTGATTAAGAAATTAASATTGATCACCATTTCAAATTT 47
Db 1 CCACGGCTTGATTAAGAAATTAAGTTGATCACCATTTCAAATTT 47
```

## RESULT 2

```
ADD69391
ID ADD69391 standard; DNA; 240823 BP.
```

```
XX ADD69391;
```

```
XX 15-JAN-2004 (first entry)
```

```
XX Human PG-3 DNA - SEQ ID 127.
```

```
XX fusion protein; fibrinogen-like; coiled-like domain;
```

```
XX angiotensin-related factor; ARF; angiotensin; Ang-1; Ang-2; Ang-3;
```

```
XX Ang-4; Ang-2X; vulnery; antiinflammatory; vasotropic; necrosis;
```

```
XX ischaemia; inflammation; wound healing; CCD; FLD; human; PG-3; de.
```

```
XX Homo sapiens.
```

```
XX WO2003048185-A2.
```

```
XX 12-JUN-2003.
```

```
XX 21-NOV-2002; 2002WO-US037660.
```

```
XX 30-NOV-2001; 2001US-0334488P.
```

```
XX (GENV-) GENVEC INC.
```

```
XX Kessler PD, Kovsed I;
```

```
XX WPI; 2003-513736/48.
```

```
XX New fusion protein comprising a fibrinogen-like or coiled-like domain,
XX useful for preparing a composition for treating necrosis, ischaemia or
XX inflammation, or for promoting wound healing.
```

```
XX Disclosure; SEQ ID NO 127; 340pp; English.
```

```
XX The invention relates to a novel fusion protein comprising a fibrinogen-
XX like domain (FLD) or coiled-coil domain (CCD). The domain may be
```

```
XX identical or homologous to that of an angiotensin-related factor (ARF),
```

```
XX examples of which include Ang (angiotensin)-1, Ang-2, Ang-3, Ang-4 and
```

```
XX Ang-2X. The molecules of the invention demonstrate vulnery,
```

```
XX antiinflammatory and vasotropic activities whilst the fusion protein may
```

```
XX be useful for preparing a composition for treating necrosis, ischaemia or
```

```
XX inflammation, as well as for promoting wound healing. The current
```

```
XX sequence is that of the human PG-3 DNA of the invention.
```

```
XX Sequence 240823 BP; 66383 A; 50394 C; 51548 G; 72418 T; 0 U; 80 Other;
```

```
XX Query Match 99.1%; Score 46.6; DB 10; Length 240823;
```

```
XX Best Local Similarity 97.9%; Pred. No. 5.2e-06;
```

```
XX Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
XX 1 CCACGGCTTGATTAAGAAATTAASATTGATCACCATTTCAAATTT 47
XX Db 83762 CCACGGCTTGATTAAGAAATTAAGTTGATCACCATTTCAAATTT 83808
```

```
XX RESULT 3
XX AAF24497
XX AAF24497 standard; cDNA; 240825 BP.
```

```
XX 23-MAY-2001 (first entry)
```



```
FT primer_bind complement(42233..42251)
FT /tag= bh
FT primer_bind 42525..42543
FT /tag= bi
FT exon 50436..50545
FT /tag= bj
FT primer_bind 67289..67309
FT /label= G
FT /tag= bk
FT primer_bind 67456..67474
FT /tag= bl
FT misc_binding 67463..67487
FT /tag= bm
FT /note= "binds probe"
FT replace(67475,G)
FT /tag= bn
FT primer_bind complement(67476..67494)
FT /tag= bo
FT primer_bind 67724..67741
FT /tag= bp
FT primer_bind 69182..69200
FT /tag= bq
FT primer_bind 69502..69520
FT /tag= br
FT misc_binding 69509..69533
FT /tag= bs
FT /note= "binds probe"
FT replace(69521,G)
FT /tag= bt
FT primer_bind complement(69522..69540)
FT /tag= bu
FT primer_bind 69609..69626
FT /tag= bv
FT primer_bind 72698..72715
FT /tag= bw
FT primer_bind 72819..72837
FT /tag= bx
FT misc_binding 72826..72850
FT /tag= by
FT /note= "binds probe"
FT replace(72838,T)
FT /tag= bz
FT primer_bind complement(72839..72857)
FT /tag= ca
FT exon 72881..72918
FT /tag= cb
FT /label= H
FT primer_bind 73099..73117
FT /tag= cc

Query Match 99.1%; Score 46.6; DB 4; Length 240825;
Best Local Similarity 97.9%; Pred. No. 5.2e-06;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACAGGCTGATTAGAAATTAATTTGATCCACATTTTCAATTTT 47
DB 83764 CCACAGGCTGATTAGAAATTAATTTGATCCACATTTTCAATTTT 83810

RESULT 4
ABO81802
ID ABO81802 standard; DNA; 240825 BP.
XX ABO81802;
XX 14-NOV-2002 (first entry)
XX Human PG-3 gene SEQ ID NO:1.
XX Human; PG-3; chromosome 8; chromosome 8p23; polymorphic; SNP;
KM single nucleotide polymorphism; biallelic marker; DNA repair;
KM recombination; cell cycle control; gene; ds.
XX
```

---

```
OS Homo sapiens.
XX Key
FH 5'UTR
FT primer_bind 1823..1840
FT /tag= a
FT primer_bind 1980..1998
FT /tag= d
FT misc_binding 1987..2011
FT /tag= nh
FT /note= "probe"
FT 1999
FT /tag= ad
FT /standard name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base G or C; the nucleotide is given
FT as an IUPAC ambiguity code in the specification"
FT complement(2000..2018)
FT /tag= dh
FT 2001..2079
FT /tag= b
FT /label= A
FT 2080..4626
FT /tag= c
FT /number= 1
FT complement(2108..2125)
FT /tag= di
FT 4459..4577
FT /tag= dj
FT primer_bind 4582..4600
FT /tag= dk
FT 4589..4613
FT /tag= ni
FT /note= "probe"
FT 4601
FT /tag= ae
FT /standard name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base A or G; the nucleotide is given
FT as an IUPAC ambiguity code in the specification"
FT complement(4602..4620)
FT /tag= dl
FT 4627..4718
FT /tag= d
FT /label= B
FT 4719..10114
FT /tag= e
FT /number= 2
FT complement(4891..4908)
FT /tag= dm
FT 10007..10025
FT /tag= dn
FT 10115..10233
FT /tag= f
FT /label= C
FT 10209..10227
FT /tag= do
FT 10216..10240
FT /tag= nj
FT /note= "probe"
FT 10228
FT /tag= af
FT /standard name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base G or T; the nucleotide is given
FT as an IUPAC ambiguity code in the specification"
FT complement(10229..10247)
FT /tag= dp
FT 10234..26809
FT /tag= g
FT /number= 3
FT 10267..10285
FT /tag= dq
FT 10274..10298
FT /tag= nk
```

```

FT /note= "probe"
FT misc_feature 10286
FT /tag= ag
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base G or T; the nucleotide is given
FT as an tUPAC ambiguity code in the specification"
FT primer_bind 10287. .10305)
FT /tag= dr
FT 10370
FT /tag= ah
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /note= "insertion of G; the nucleotide is given as an
FT tUPAC ambiguity code in the specification"
FT primer_bind 10411. .10430)
FT /tag= dg
FT 26810. .26897
FT /tag= h
FT /label= D
FT 26898. .31356
FT /tag= i
FT /number= 4
FT 31357. .31471
FT /tag= j
FT /label= E
FT 31472. .34260
FT /tag= k
FT /number= 5
FT 34261. .34404
FT /tag= l
FT /label= F
FT 34405. .37376
FT /tag= m
FT /number= 6
FT 37377. .37466
FT /tag= n
FT /label= S
FT 37467. .39703
FT /tag= o
FT /number= 7
FT 39556. .39574
FT /tag= dc
FT 39704. .40858
FT /tag= p
FT /label= T
FT 39877. .39896
FT /tag= du
FT 39925. .39943
FT /tag= dv
FT 39933. .39956
FT /tag= nl
FT /note= "probe"
FT 39944
FT /tag= ai
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base G or T; the nucleotide is given
FT as an tUPAC ambiguity code in the specification"
FT primer_bind 39945. .39963)
FT /tag= dw
FT /complement(39953. .39970)
FT /tag= dx
FT 39954. .39972
FT /tag= dy
FT 39961. .39985
FT /tag= nm
FT /note= "probe"
FT 39973
FT /tag= aj
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base G or C; the nucleotide is given
FT as an tUPAC ambiguity code in the specification"
FT primer_bind 39974. .39992)
FT /tag= dz
FT /complement(40242. .40259)

```

```

FT /tag= ea
FT intron 40859. .50435
FT /tag= q
FT /number= 8
FT primer_bind 41137. .41154
FT /tag= eb
FT primer_bind 41366. .41384
FT /tag= ec
FT 41373. .41397
FT /tag= nm
FT /note= "probe"
FT 41385. .41403
FT /tag= ed
FT 41385
FT /tag= ak
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base A or C; the nucleotide is given
FT as an tUPAC ambiguity code in the specification"
FT primer_bind 41386. .41404)
FT /tag= ee
FT 41392. .41416
FT /tag= no
FT /note= "probe"
FT 41404
FT /tag= al
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base A or C; the nucleotide is given
FT as an tUPAC ambiguity code in the specification"
FT primer_bind 41405. .41423)
FT /tag= ef
FT /complement(41564. .41581)
FT /tag= eg
FT 42122. .42141
FT /tag= eh
FT 42213. .42231
FT /tag= ei
FT 42220. .42244
FT /tag= np
FT /note= "probe"
FT 42232
FT /tag= am
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT misc_feature 42232

Query Match 99.1%; Score 46.6; DB 6; Length 240825;
Best Local Similarity 97.9%; Pred. No. 5.2e-06;
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCACAGGCTTGATGAATAAATTAATTTGATCACCAATTTCAATTTT 47
Db 83764 CCACAGGCTTGATGAATAAATTAAGTTGATCACCAATTTCAATTTT 83810

RESULT 5
ABK31511/c
ID ABK31511 standard; DNA; 47108 BP.
XX
AC ABK31511;
XX
DT 23-APR-2002 (first entry)
XX
DE Signal transduction associated gene modified complementary DNA #177.
XX
KW Human; signal transduction associated gene; cytosine methylation state;
KW Cpg island; signal transduction associated disease; solid tumour; cancer;
KW antitumour; cytosstatic; mutant; de.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200200926-A2.
XX
PD 03-JAN-2002.
XX

```

PF 29-JUN-2001; 2001WO-EP007472.  
XX  
XX 30-JUN-2000; 2000DE-01032529.  
PR 01-SEP-2000; 2000DE-01043826.  
XX  
XX (EPIC-) EPIDENOMICS AG.  
XX  
XX Olek A, Piepenbrock C, Berlin K;  
XX  
XX WPI; 2002-147896/19.  
XX  
XX Oligonucleotide for diagnosis and therapy of diseases associated with  
PT signal transduction e.g. cancer, comprises chemically modified genomic  
PT sequences of genes associated with signal transduction.  
XX  
XX Claim 1; SEQ ID NO 354; 24pp; English.  
XX  
XX The present invention relates to chemically modified DNA sequences of  
CC signal transduction associated genes. The DNA sequences are chemically  
CC modified using a solution of bisulphite, hydrogen sulphite or disulphite.  
CC Also disclosed are oligonucleotides and/or PNA oligomers for detecting  
CC the cytosine methylation state (CpG islands) of these genes, and a method  
CC for the diagnosis and/or therapy of genetic and epigenetic parameters of  
CC genes associated with signal transduction. The genomic DNA can be  
CC obtained from cells or cellular components which contain DNA, e.g. cell  
CC lines, biopsies, blood, sputum, stool, urine, cerebral-spinal fluid,  
CC tissue embedded in paraffin such as tissue from eyes, intestine, kidney,  
CC brain, heart, prostate, lung, breast or liver, histologic object slides,  
CC and all their possible combinations. The sequences of the invention are  
CC useful for the diagnosis and therapy of diseases associated with signal  
CC transduction e.g. solid tumours and cancer. ABK3158-ABK31545 represent  
CC chemically pretreated genomic DNA sequences of different genes associated  
CC with signal transduction, or their complementary sequences. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC European Patent Office  
XX  
XX Sequence 47108 BP; 15349 A; 429 C; 8736 G; 22594 T; 0 U; 0 Other;  
SQ  
XX  
XX Query Match 51.9%; Score 24.4; DB 6; Length 47108;  
Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
Matches 28; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
OY 12 ATTGAAATTAASSTTGATCACCATTTCGAATTTT 47  
DB 25070 ATTAATAAACTTTAATTAATTTTAAATTTT 25035  
RESULT 6  
ADA32589/C  
ID ADA32589 standard; DNA; 993 BP.  
XX  
XX ADA32589;  
AC  
XX 20-NOV-2003 (first entry)  
DT  
XX  
XX DNA encoding *Acinetobacter baumannii* protein #3876.  
DE  
XX  
XX ds; gene; *Acinetobacter baumannii*; bacterial disease; antibacterial;  
KW vaccine; plant biocontrol agent.  
KM  
XX  
XX *Acinetobacter baumannii*.  
OS  
XX  
XX US6562958-B1.  
PN  
XX  
XX 13-MAY-2003.  
PD  
XX  
XX 04-JUN-1999; 99US-00328352.  
PF  
XX  
XX 09-JUN-1998; 98US-0088701P.  
PR  
XX  
XX (GENO-) GENOME THERAPEUTICS CORP.  
PA  
XX

PI Breton G, Bush D;  
XX  
XX WPI; 2003-576092/54.  
DR  
XX P-PSDB; ADA36715.  
DR  
XX  
XX New *Acinetobacter baumannii* proteins and nucleic acids, useful as reagents  
PT for diagnosing a bacterial disease, as components of antibacterial  
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for  
PT plants.  
XX  
XX Example; SEQ ID NO 3876; 328pp; English.  
PS  
XX  
XX The invention relates to isolated *Acinetobacter baumannii* nucleic acids.  
CC The A. *baumannii* nucleic acids and polypeptides are useful as reagents  
CC for diagnosing a bacterial disease, as components of antibacterial  
CC vaccines, as targets for antibacterial drugs, to detect the presence of  
CC A. *baumannii* and other *Acinetobacter* species in a sample, in screening  
CC compounds for the ability to interfere with the A. *baumannii* life cycle  
CC or to inhibit A. *baumannii* infection, and as biocontrol agents for  
CC plants. The present sequence represents DNA encoding an A. *baumannii*  
CC protein.  
XX  
XX Sequence 993 BP; 309 A; 181 C; 203 G; 300 T; 0 U; 0 Other;  
SQ  
XX  
XX Query Match 51.5%; Score 24.2; DB 9; Length 993;  
Best Local Similarity 68.1%; Pred. No. 1.3e+02;  
Matches 32; Conservative 1; Mismatches 14; Indels 0; Gaps 0;  
OY 1 CCACGCGCTGATTGAAATTAASSTTGATCACCATTTCGAATTTT 47  
DB 205 CACACGTTTACTGACAGATTAATGCTGATTAATTTTGACACTTT 159  
RESULT 7  
ABL07356/C  
ID ABL07356 standard; cDNA; 4937 BP.  
XX  
XX ABL07356;  
AC  
XX  
XX 26-MAR-2002 (first entry)  
DT  
XX  
XX *Drosophila melanogaster* expressed polynucleotide SEQ ID NO 16550.  
DE  
XX  
XX *Drosophila*; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ss.  
KM  
XX  
XX *Drosophila melanogaster*.  
OS  
XX  
XX WO200171042-A2.  
PN  
XX  
XX 27-SEP-2001.  
PD  
XX  
XX 23-MAR-2001; 2001WO-US009231.  
PF  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
PR  
XX  
XX 11-JUL-2000; 2000US-00614150.  
PR  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI  
XX WPI; 2001-656860/75.  
DR  
XX P-PSDB; ABB63253.  
DR  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from *Drosophila* and for elucidating cell signalling and cell-cell  
PT interactions.  
XX  
XX Claim 1; SEQ ID NO 16550; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is  
CC useful in developmental biology and in elucidating cell signalling and



CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB15176-AB130511), expressed DNA  
CC sequences (AB101840-AB15175) and the encoded proteins (AAB57737-  
CC AAB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

SO Sequence 4937 BP; 1248 A; 1272 C; 1143 G; 1274 T; 0 U; 0 Other;

Query Match 51.5%; Score 24.2; DB 4; Length 4937;

Best Local Similarity 68.1%; Pred. No. 1.5e+02;  
Matches 32; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

OY 1 CCACAGCGTTGATTAGAAATTAATTAATTTGATCCACATTTCATTAATT 47  
DB 2273 CAACAGTGGCTATGAGTAAATTAAGTTTAAACACACTTTCATTAATT 2227

RESULT 8

ABL32951/C  
ID ABL32951 standard; DNA; 5173 BP.

XX ABL32951;

XX 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 924.

XX Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antiasthmatic; cytosine; nootropic;  
KW neuroprotective; anti-HIV; anticoagulant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
KW ds.

XX Homo sapiens.

XX WO200200928-A2.

XX 03-JAN-2002.

PD 02-JUL-2001; 2001WO-EP007537.

PR 30-JUN-2000; 2000DE-01032529.

PR 01-SEP-2000; 2000DE-01043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

DR WPI; 2002-130909/17.

PT Nucleic acid comprising fragment of chemically modified gene, useful for  
PT diagnosis and treatment of diseases associated with abnormal cytosine  
PT methylation.

PS Claim 1; SEQ ID NO 924; 32pp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention

SO Sequence 5173 BP; 1274 A; 129 C; 1256 G; 2508 T; 0 U; 6 Other;

Query Match 51.5%; Score 24.2; DB 6; Length 5173;

Best Local Similarity 71.1%; Pred. No. 1.5e+02;  
Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 1 CCACAGCGTTGATTAGAAATTAATTAATTTGATCCACATTTCATTAATT 45

DB 3820 CCATATCCCTTAATTAATAATTAATTAATTTCAATCACTATTCCTAATT 3776

RESULT 9

AAS36498  
ID AAS36498 standard; DNA; 27681 BP.

XX AAS36498;

XX 17-DEC-2001 (first entry)

DE Human cardiovascular system antigen genomic DNA SEQ ID No 1998.

XX Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;  
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;  
KW antirheumatic; antiproliferative; cytosolic; cardiant; neuroprotective;  
KW cerebroprotective; nootropic; antibacterial; vitruide; fungicide; cancer;  
KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; tissue regeneration;  
KW anti-infertility.

XX Homo sapiens.

XX WO200155321-A2.

XX 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US001340.

PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 17-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 14-JUL-2000; 2000US-0217496P.

PR 26-JUL-2000; 2000US-0220963P.

PR 16-AUG-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225477P.

PR 14-AUG-2000; 2000US-0225575P.

PR 14-AUG-2000; 2000US-0225585P.

PR 14-AUG-2000; 2000US-0225595P.

PR 18-AUG-2000; 2000US-0226799P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226688P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.



RESULT 10  
AAS36497  
ID AAS36497 standard; DNA: 27681 BP.  
XX AAS36497;  
DT 17-DEC-2001 (first entry)  
XX  
DE Human cardiovascular system antigen genomic DNA SEQ ID No 1997.  
XX  
KW Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;  
KW chicken; sheep; immunosuppressive; antichrictic; vasotropic; dog;  
KW antithematic; antiproliferative; cytostatic; cardiant; neuroprotective;  
KW cerebroprotective; noctropic; antibacterial; virucide; fungicide; cancer;  
KW opthalmological; vulnerary; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; tissue regeneration;  
KW anti-fertility.  
XX  
OS Homo sapiens.  
XX  
PN WO200155321-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001340.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-022679P.  
PR 22-AUG-2000; 2000US-022681P.  
PR 22-AUG-2000; 2000US-022688P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231988P.  
PR 14-SEP-2000; 2000US-023297P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-023423P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234597P.  
PR 25-SEP-2000; 2000US-0234598P.  
PR 25-SEP-2000; 2000US-0234584P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.



PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234597P.  
PR 25-SEP-2000; 2000US-0234598P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251930P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Barash SC, Ruben SM,  
XX WPI, 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX  
XX Disclosure; SEQ ID NO 40655; 3071pp + Sequence Listing; English.  
PS  
XX  
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention  
XX  
SQ Sequence 27681 BP; 7624 A; 6057 C; 5942 G; 8058 T; 0 U; 0 Other;  
Query Match 50.2%; Score 23.6; DB 4; Length 27681;  
Best Local Similarity 76.3%; Pred. No. 2.9e+02;  
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 8 CTTGATTAGAAATTAATTTGATCACCATTTCATTT 45  
DB 25518 CTTGATGAGAGTAATTTAGAACATTAATTTTCATTT 25555  
|||||  
RESULT 12  
ADE47191  
ID ADE47191 standard; DNA; 27681 BP.  
XX  
AC ADE47191;  
XX  
XX 29-JAN-2004 (first entry)  
DT  
XX  
DE Human cardiovascular system related genomic DNA #757.  
XX  
XX Human; cardiovascular system related polypeptide; cancer;  
KW proliferative disorder; foetal abnormality; developmental abnormality;  
KW haematopoietic disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
KW inflammation; allergy; neurological disorder; Alzheimer's disease;  
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;  
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;  
KW cardiovascular disorder; angioedema disorder; kidney disorder;  
KW gastrointestinal disorder; pregnancy-related disorder;  
KW endocrine disorder; gene; de.  
XX  
OS Homo sapiens.  
XX  
PN US2003059908-A1.  
XX  
XX 27-MAR-2003.  
PD  
XX  
PF 07-MAR-2002; 2002US-00091504.

XX  
PR 31-JAN-2000; 2000US-0179065P-  
PR 04-FEB-2000; 2000US-0180628P-  
PR 24-FEB-2000; 2000US-0184664P-  
PR 02-MAR-2000; 2000US-0186350P-  
PR 16-MAR-2000; 2000US-0189874P-  
PR 17-MAR-2000; 2000US-0190076P-  
PR 18-APR-2000; 2000US-0198123P-  
PR 19-MAY-2000; 2000US-0205515P-  
PR 07-JUN-2000; 2000US-0203467P-  
PR 28-JUN-2000; 2000US-0214886P-  
PR 30-JUN-2000; 2000US-0215135P-  
PR 07-JUL-2000; 2000US-0216647P-  
PR 07-JUL-2000; 2000US-0216880P-  
PR 11-JUL-2000; 2000US-0217487P-  
PR 14-JUL-2000; 2000US-0217496P-  
PR 26-JUL-2000; 2000US-0218290P-  
PR 26-JUL-2000; 2000US-0220963P-  
PR 14-AUG-2000; 2000US-0220964P-  
PR 14-AUG-2000; 2000US-0224518P-  
PR 14-AUG-2000; 2000US-0225213P-  
PR 14-AUG-2000; 2000US-0225214P-  
PR 14-AUG-2000; 2000US-0225266P-  
PR 14-AUG-2000; 2000US-0225267P-  
PR 14-AUG-2000; 2000US-0225268P-  
PR 14-AUG-2000; 2000US-0225270P-  
PR 14-AUG-2000; 2000US-0225447P-  
PR 14-AUG-2000; 2000US-0225757P-  
PR 14-AUG-2000; 2000US-0225758P-  
PR 14-AUG-2000; 2000US-0225759P-  
PR 18-AUG-2000; 2000US-0226279P-  
PR 22-AUG-2000; 2000US-0226681P-  
PR 22-AUG-2000; 2000US-0226688P-  
PR 22-AUG-2000; 2000US-0227182P-  
PR 23-AUG-2000; 2000US-0227009P-  
PR 30-AUG-2000; 2000US-0228924P-  
PR 01-SEP-2000; 2000US-0229287P-  
PR 01-SEP-2000; 2000US-0229343P-  
PR 01-SEP-2000; 2000US-0229344P-  
PR 01-SEP-2000; 2000US-0229345P-  
PR 03-SEP-2000; 2000US-0229309P-  
PR 05-SEP-2000; 2000US-0229513P-  
PR 06-SEP-2000; 2000US-0230437P-  
PR 06-SEP-2000; 2000US-0230438P-  
PR 08-SEP-2000; 2000US-0231242P-  
PR 08-SEP-2000; 2000US-0231243P-  
PR 08-SEP-2000; 2000US-0231244P-  
PR 08-SEP-2000; 2000US-0231413P-  
PR 08-SEP-2000; 2000US-0231414P-  
PR 08-SEP-2000; 2000US-0233080P-  
PR 08-SEP-2000; 2000US-0233081P-  
PR 12-SEP-2000; 2000US-0233968P-  
PR 14-SEP-2000; 2000US-0233397P-  
PR 14-SEP-2000; 2000US-0233398P-  
PR 14-SEP-2000; 2000US-0233399P-  
PR 14-SEP-2000; 2000US-0233400P-  
PR 14-SEP-2000; 2000US-0233401P-  
PR 14-SEP-2000; 2000US-0233063P-  
PR 14-SEP-2000; 2000US-0233064P-  
PR 14-SEP-2000; 2000US-0233065P-  
PR 21-SEP-2000; 2000US-0234223P-  
PR 21-SEP-2000; 2000US-0234274P-  
PR 25-SEP-2000; 2000US-0234997P-  
PR 25-SEP-2000; 2000US-0234998P-  
PR 26-SEP-2000; 2000US-0234984P-  
PR 27-SEP-2000; 2000US-0235834P-  
PR 27-SEP-2000; 2000US-0235835P-  
PR 29-SEP-2000; 2000US-0236327P-  
PR 29-SEP-2000; 2000US-0236367P-  
PR 29-SEP-2000; 2000US-0236368P-  
PR 29-SEP-2000; 2000US-0236369P-  
PR 29-SEP-2000; 2000US-0236370P-  
PT

PR 02-OCT-2000; 2000US-0236802P-  
PR 02-OCT-2000; 2000US-0237037P-  
PR 02-OCT-2000; 2000US-0237038P-  
PR 02-OCT-2000; 2000US-0237039P-  
PR 02-OCT-2000; 2000US-0237040P-  
PR 13-OCT-2000; 2000US-0239935P-  
PR 13-OCT-2000; 2000US-0239937P-  
PR 20-OCT-2000; 2000US-0240960P-  
PR 20-OCT-2000; 2000US-0241221P-  
PR 20-OCT-2000; 2000US-0241785P-  
PR 20-OCT-2000; 2000US-0241786P-  
PR 20-OCT-2000; 2000US-0241787P-  
PR 20-OCT-2000; 2000US-0241808P-  
PR 20-OCT-2000; 2000US-0241809P-  
PR 20-OCT-2000; 2000US-0241826P-  
PR 01-NOV-2000; 2000US-0244617P-  
PR 08-NOV-2000; 2000US-0246474P-  
PR 08-NOV-2000; 2000US-0246475P-  
PR 08-NOV-2000; 2000US-0246476P-  
PR 08-NOV-2000; 2000US-0246477P-  
PR 08-NOV-2000; 2000US-0246478P-  
PR 08-NOV-2000; 2000US-0246523P-  
PR 08-NOV-2000; 2000US-0246524P-  
PR 08-NOV-2000; 2000US-0246525P-  
PR 08-NOV-2000; 2000US-0246526P-  
PR 08-NOV-2000; 2000US-0246527P-  
PR 08-NOV-2000; 2000US-0246528P-  
PR 08-NOV-2000; 2000US-0246532P-  
PR 08-NOV-2000; 2000US-0246609P-  
PR 08-NOV-2000; 2000US-0246610P-  
PR 08-NOV-2000; 2000US-0246611P-  
PR 08-NOV-2000; 2000US-0246613P-  
PR 17-NOV-2000; 2000US-0249207P-  
PR 17-NOV-2000; 2000US-0249208P-  
PR 17-NOV-2000; 2000US-0249209P-  
PR 17-NOV-2000; 2000US-0249210P-  
PR 17-NOV-2000; 2000US-0249211P-  
PR 17-NOV-2000; 2000US-0249212P-  
PR 17-NOV-2000; 2000US-0249213P-  
PR 17-NOV-2000; 2000US-0249214P-  
PR 17-NOV-2000; 2000US-0249215P-  
PR 17-NOV-2000; 2000US-0249216P-  
PR 17-NOV-2000; 2000US-0249217P-  
PR 17-NOV-2000; 2000US-0249218P-  
PR 17-NOV-2000; 2000US-0249299P-  
PR 17-NOV-2000; 2000US-0249300P-  
PR 01-DEC-2000; 2000US-0250160P-  
PR 01-DEC-2000; 2000US-0250391P-  
PR 05-DEC-2000; 2000US-0251030P-  
PR 05-DEC-2000; 2000US-0251988P-  
PR 05-DEC-2000; 2000US-0256719P-  
PR 06-DEC-2000; 2000US-0251479P-  
PR 08-DEC-2000; 2000US-0251865P-  
PR 08-DEC-2000; 2000US-0251868P-  
PR 08-DEC-2000; 2000US-0251869P-  
PR 08-DEC-2000; 2000US-0251989P-  
PR 08-DEC-2000; 2000US-0251990P-  
PR 11-DEC-2000; 2000US-0254097P-  
PR 05-JAN-2001; 2001US-0259678P-  
PR 17-JAN-2001; 2001US-00764869.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Ruben SM, Barash SC;  
XX WPI; 2003-743766/70.  
XX  
PT New cardiovascular system related polynucleotides and polypeptides,

PT useful for preventing, treating, or ameliorating a medical condition,  
PT such as cancer of cardiovascular tissues and cancer metastases.  
XX  
PS Claim 1, SEQ ID NO 1997, 262pp, English.  
XX  
XX The invention relates to human cardiovascular system related polypeptides  
CC and the polynucleotides encoding them. The polypeptides, polynucleotides  
CC and antibodies to the polypeptides are useful for diagnosing a  
CC pathological condition or a susceptibility to a pathological condition,  
CC for preventing, treating, or ameliorating a medical condition, such as  
CC cancer of cardiovascular system tissues, proliferative disorders, foetal  
CC and developmental abnormalities, haematopoietic disorders, diseases of  
CC the immune system, AIDS, autoimmune diseases (e.g., rheumatoid  
CC arthritis), inflammation, allergies, neurological disorders (e.g.,  
CC Alzheimer's disease, Parkinson's disease), cognitive disorders,  
CC schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis,  
CC diabetes, atherosclerosis, cardiovascular disorders, angiogenic  
CC disorders, kidney disorders, gastrointestinal disorders, pregnancy-  
CC related disorders, endocrine disorders and infections. The nucleic acids  
CC are also useful for chromosome identification, radiation hybrid mapping  
CC or long-range restriction mapping. The polypeptides and polynucleotides  
CC may also be used as food additives or preservatives to increase or  
CC decrease storage capabilities, fat content or other nutritional  
CC components. This sequence represents human cardiovascular system related  
CC genomic DNA of the invention.  
XX  
SQ Sequence 27681 BP; 7624 A; 6057 C; 5942 G; 8058 T; 0 U; 0 Other;  
Query Match 50.2%; Score 23.6; DB 10; Length 27681;  
Best Local Similarity 76.3%; Pred. No. 2.9e+02;  
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
Qy 8 CTTGATTAGAAATTAATTTGATCACCATTTCAAATT 45  
Db 25518 CTTGATGAGAGTAATTTAGAACATTATTTTCATATT 25555  
RESULT 13  
ADE47192  
ID ADE47192 standard; DNA; 27681 BP.  
XX  
AC ADE47192;  
XX  
XX 29-JAN-2004 (first entry)  
XX  
XX Human cardiovascular system related genomic DNA #758.  
XX  
DE Human cardiovascular system related polypeptide; cancer;  
XX proliferative disorder; foetal abnormality; developmental abnormality;  
KW haematopoietic disorder; AIDS; autoimmune disease; Rheumatoid arthritis;  
KW inflammation; allergy; neurological disorder; Alzheimer's disease;  
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;  
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;  
KW cardiovascular disorder; angiogenic disorder; kidney disorder;  
KW gastrointestinal disorder; pregnancy-related disorder;  
KW endocrine disorder; gene; ds.  
XX  
XX Homo sapiens.  
OS  
XX  
PN US2003059908-A1.  
XX  
XX 27-MAR-2003.  
PD  
XX  
XX 07-MAR-2002; 2002US-00091504.  
PF  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231245P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-02331988P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234497P.  
PR 25-SEP-2000; 2000US-0234988P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.



PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR 17-JAN-2001; 2001US-0076486P.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Ruden SM, Barash SC;  
XX WPI, 2003-743766/70.  
XX  
XX  
PT New cardiovascular system related polynucleotides and polypeptides,  
PT useful for preventing, treating, or ameliorating a medical condition,  
PT such as cancer of cardiovascular tissues and cancer metastases.  
XX  
XX  
PS Claim 1; SEQ ID NO 1998; 262zp; English.  
XX  
XX  
CC The invention relates to human cardiovascular system related polypeptides  
CC and the polynucleotides encoding them. The polypeptides, polynucleotides  
CC and antibodies to the polypeptides are useful for diagnosing a  
CC pathological condition or a susceptibility to a pathological condition.

CC for preventing, treating, or ameliorating a medical condition, such as  
CC cancer of cardiovascular system tissues, proliferative disorders, foetal  
CC and developmental abnormalities, haematopoietic disorders, diseases of  
CC the immune system, AIDS, autoimmune diseases (e.g., rheumatoid  
CC arthritis), inflammation, allergies, neurological disorders (e.g.,  
CC Alzheimer's disease, Parkinson's disease), cognitive disorders,  
CC schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis,  
CC diabetes, atherosclerosis, cardiovascular disorders, angiogenic  
CC disorders, kidney disorders, gastrointestinal disorders, pregnancy-  
CC related disorders, endocrine disorders and infections. The nucleic acids  
CC are also useful for chromosome identification, radiation hybrid mapping  
CC or long-range restriction mapping. The polypeptides and polynucleotides  
CC may also be used as food additives or preservatives to increase or  
CC decrease storage capabilities, fat content or other nutritional  
CC components. This sequence represents human cardiovascular system related  
CC genomic DNA of the invention.  
XX  
SQ Sequence 27681 BP; 7624 A; 6057 C; 5942 G; 8058 T; 0 U; 0 Other;  
Query Match 50.2%; Score 23.6; DB 10; Length 27681;  
Best Local Similarity 76.3%; Pred. No. 2.9e+02;  
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
Qy 8 CTTGATTGGAATTAATTTGATCCATTTCGAATT 45  
Db 25518 CTTGATGAGAGTAATTAGAACATTTTCATATT 25555  
RESULT 14  
ADJ08610  
ADJ08610 standard; DNA; 27681 BP.  
XX  
AC ADJ08610;  
XX  
DT 04-NOV-2004 (first entry)  
XX  
DE Human cardiovascular system associated polypeptide-related DNA SeqID1998.  
XX  
KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;  
KW breast neoplasms; liver neoplasm; cardiovascular disorder;  
KW cardiac arrest; cerebrovascular disorder; cerebral ischaemia;  
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;  
KW ocular disorder; corneal infection; wound healing;  
KW epithelial cell proliferation; skin aging; sunburn;  
KW organ transplantation; cell culture; tissue regeneration; chemotaxis;  
KW food additive; preservative; cardiovascular system associated antigen;  
KW nuclear factor kappaB; NFkappaB; promoter element; human; ds.  
XX  
OS Homo sapiens.  
XX  
PN US2004005575-A1.  
XX  
PD 08-JAN-2004.  
XX  
PF 26-AUG-2002; 2002US-00227577.  
XX  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-023142P.  
PR 08-SEP-2000; 2000US-023143P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234977P.  
PR 25-SEP-2000; 2000US-0234988P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235634P.  
PR 27-SEP-2000; 2000US-0235636P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-024121P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0251989P.  
PR 06-DEC-2000; 2000US-0251792P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR 17-JAN-2001; 2001US-0276486P.  
PR 07-MAR-2002; 2002US-00091504.  
  
(HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX  
PI Rosen CA, Ruben SM, Barash SC,  
XX  
XX WPI; 2004-081713/08.  
DR  
XX  
XX  
PT New cardiovascular system-related nucleic acid molecule, useful for  
PT diagnosing, preventing or treating diseases of the cardiovascular system,  
PT and in chromosome mapping, drug screening or in pharmacogenomics.  
XX  
XX  
PS Disclosure; SEQ ID NO 1998; 262pp; English.  
  
XX The invention relates to an isolated nucleic acid molecule encoding a  
CC human cardiovascular system associated polypeptide (or antigen), or its  
CC fragment. Also included recombinant vectors, recombinant host cells, an  
CC isolated human cardiovascular system associated polypeptide (including  
CC its fragment, allelic variant, species homologue or epitope), an isolated  
CC antibody that binds specifically to a human cardiovascular system  
CC associated polypeptide, diagnosing a pathological condition or  
CC susceptibility to a pathological condition (comprising determining the  
CC presence or absence of a mutation in human cardiovascular system  
CC associated nucleic acid and diagnosing a condition based on the presence  
CC or absence of the mutation), identifying a binding partner to human  
CC cardiovascular system associated polypeptides, the gene corresponding to

CC the human cardiovascular system associated cDNA sequence and identifying  
CC an activity in a biological assay comprising expressing the human  
CC cardiovascular system associated cDNA in a cell, isolating the  
CC supernatant, detecting an activity in a biological assay and identifying  
CC the protein in the supernatant having the activity. The human  
CC cardiovascular system associated nucleic acids and polypeptides are used  
CC to prevent, treat or ameliorate a medical condition (for example in  
CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep), for  
CC example autoimmune diseases such as rheumatoid arthritis,  
CC hyperproliferative disorders, for example neoplasms of the breast or  
CC liver, cardiovascular disorders, for example cardiac arrest,

Query Match 50.2%; Score 23.6; DB 13; Length 27681;  
Best Local Similarity 76.3%; Pred. No. 2.9e+02;  
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 8 CTTGATTGAATTAATTTGATCCACATTTTCAATT 45  
Db 25518 CTTGATGAGAAAGTAAATTTAGAACATTAATTTCAATT 25555

## RESULT 15

ADJ08609 standard; DNA; 27681 BP.

ADJ08609;

04-NOV-2004 (first entry)

Human cardiovascular system associated polypeptide-related DNA SeqID1997.

KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;  
KW breast neoplasms; liver neoplasm; cardiovascular disorder;  
KW cardiac arrest; cerebrovascular disorder; cerebral ischaemia;  
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;  
KW ocular disorder; corneal infection; skin aging; sunburn;  
KW epithelial cell proliferation; cell culture; tissue regeneration; chemotaxis;  
KW organ transplantation; cell culture; tissue regeneration; chemotaxis;  
KW food additive; preservative; cardiovascular system associated antigen;  
KW nuclear factor kappaB; NFkappaB; promoter element; human; ds.

XX Homo sapiens.

XX US2004005575-A1.

PD 08-JAN-2004.

PF 26-AUG-2002; 2002US-00227577.

XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225265P.  
PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227099P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0228287P.  
PR 01-SEP-2000; 2000US-0228343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231245P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0232402P.  
PR 14-SEP-2000; 2000US-0232403P.  
PR 14-SEP-2000; 2000US-0232404P.  
PR 14-SEP-2000; 2000US-0232405P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234977P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0234984P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0244674P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR 17-JAN-2001; 2001US-00764869.  
PR 07-MAR-2002; 2002US-00091504.  
  
(HUMA-) HUMAN GENOME SCI INC.  
XX  
PA  
PI Rosen CA, Ruben SM, Barash SC;  
XX  
DR WPI; 2004-081713/08.  
XX  
PT New cardiovascular system-related nucleic acid molecule, useful for  
XX diagnosing, preventing or treating diseases of the cardiovascular system,  
XX and in chromosome mapping, drug screening or in pharmacogenomics.  
XX  
PS Disclosure; SEQ ID NO 1997; 262bp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule encoding a  
CC human cardiovascular system associated polypeptide (or antigen), or its  
CC fragment. Also included recombinant vectors, recombinant host cells, an  
CC isolated human cardiovascular system associated polypeptide (including  
CC its fragment, allelic variant, species homologue or epitope), an isolated  
CC antibody that binds specifically to a human cardiovascular system  
CC associated polypeptide, diagnosing a pathological condition or  
CC susceptibility to a pathological condition (comprising determining the  
CC presence or absence of a mutation in human cardiovascular system  
CC associated nucleic acid and diagnosing a condition based on the presence  
CC or absence of the mutation), identifying a binding partner to human  
CC cardiovascular system associated polypeptides, the gene corresponding to  
CC the human cardiovascular system associated cDNA sequence and identifying  
CC an activity in a biological assay comprising expressing the human  
CC cardiovascular system associated cDNA in a cell, isolating the  
CC supernatant, detecting an activity in a biological assay and identifying  
CC the protein in the supernatant having the activity. The human  
CC cardiovascular system associated nucleic acids and polypeptides are used  
CC to prevent, treat or ameliorate a medical condition (for example in

CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep), for  
CC example autoimmune diseases such as rheumatoid arthritis,  
CC hyperproliferative disorders, for example neoplasms of the breast or  
CC liver, cardiovascular disorders, for example cardiac arrest.  
  
Query Match 50.2%; Score 23.6; DB 13; Length 27681;  
Best Local Similarity 76.3%; Pred. No. 2.9e+02;  
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
  
QY 8 CTTGATTAGAAATTAATTTGATCACCATTTCGAATT 45  
DB 25518 CTTGATGAGAGTAATTTAGAACATTATTTCAATT 25555  
  
RESULT 16  
AAH93984/C  
ID AAH93984 standard; cDNA; 342 BP.  
XX  
AC AAH93984;  
XX  
DT 05-OCT-2001 (first entry)  
XX  
XX Human foetal cDNA, SEQ ID NO: 513.  
DE Human foetal cDNA, SEQ ID NO: 513.  
XX  
XX Human; foetal protein; cytostatic; immunosuppressive; immunostimulant;  
XX neurotropic; neuroprotective; thrombolytic; osteoparitic; antiinflammatory;  
XX gene therapy; antisense therapy; cancer; immune disorder;  
XX growth disorder; osteoporosis; thrombolytic disorder;  
XX nervous system disorder; inflammation; expressed sequence tag; EST; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200155339-A2.  
XX  
XX 02-AUG-2001.  
PD  
XX 25-JAN-2001; 2001WO-US002723.  
XX  
XX 25-JAN-2000; 2000US-00491404.  
PR 15-SEP-2000; 2000US-00663870.  
PR 06-NOV-2000; 2000US-00707351.  
XX  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
XX Yeung G, Ford JE, Boyle BJ, Arterburn MC, Dormanac RA, Tang YT;  
PI Liu C, Asundi V, Zhou P, Wetman T;  
XX  
XX WPI; 2001-465571/50.  
DR P-PSDB; AAM06309.  
XX  
XX Novel fetal proteins useful for the treatment and diagnosis of diseases  
XX associated with dysfunction of the protein e.g. cancers, immune  
XX disorders, growth disorders, thrombolytic disorders, nervous system  
XX disorders and inflammation.  
XX  
PS Claim 1, Page 365; 715bp; English.  
XX  
XX The invention relates to novel foetal polypeptides encoded by  
CC polynucleotides comprising one of 477 sequences fully defined in the  
CC specification. The foetal polynucleotides and polypeptides are useful in  
CC the treatment and diagnosis of diseases such as cancers, immune  
CC disorders, growth disorders (e.g. osteoporosis), thrombolytic disorders,  
CC nervous system disorders and inflammation. The present sequence was  
CC assembled using an expressed sequence tag (EST) found to be expressed in  
XX human foetal tissue cDNA libraries as the seed  
XX  
XX Sequence 342 BP; 125 A; 46 C; 57 G; 114 T; 0 U; 0 Other;  
QY Query Match 49.8%; Score 23.4; DB 5; Length 342;  
Best Local Similarity 69.8%; Pred. No. 2.2e+02;  
Matches 30; Conservative 1; Mismatches 12; Indels 0; Gaps 0;  
QY 5 AGGCTTGATTAGAAATTAATTTGATCACCATTTCGAATT 47

Db 99 AGGATTGATTCYAACTTAACATTAACTACGACATTATTAATATTTT 57  
RESULT 17  
ID AAK70686/c  
XX AAK70686 standard; DNA; 34917 BP.  
XX AAK70686;  
DT 06-NOV-2001 (first entry)  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25498.  
XX  
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
OS Homo sapiens.  
PN WO200157182-A2.  
PD 09-AUG-2001.  
PF 17-JAN-2001; 2001WO-US001354.  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 11-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 30-AUG-2000; 2000US-0227099P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0233401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234972P.  
PR 25-SEP-2000; 2000US-0234984P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239395P.  
PR 13-OCT-2000; 2000US-0239397P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0244674P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246612P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.



XX	29-OCT-2001; 2001WO-GB004789.
PF	
XX	27-OCT-2000; 2000GB-00026333.
XX	24-NOV-2000; 2000GB-00028727.
PR	07-MAR-2001; 2001GB-00005640.
XX	
PA	(CHIR-) CHIRON SPA.
PA	(GENO-) INST GENOMIC RES.
XX	
P1	Telford J, Masignani V, Margarit y Rosi, Grandi G, Fraser C,
PI	Tetteijn H,
XX	
DR	WPI: 2002-352536/38.
DR	P-PSDB; ABP27063.
XX	
PT	New Streptococcus protein for the treatment or prevention of infection or
PT	disease caused by Streptococcus bacteria, such as meningitis, and for
PT	detecting a compound that binds to the protein.
XX	
XX	Claim 7; Page 3484; 4525pp; English.
PS	
XX	
CC	The invention relates to a protein (ABP25413-ABP30895) from group B
CC	streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC	(Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC	the specification. The proteins have antibacterial and anti-inflammatory
CC	activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
CC	antibodies that bind (I) are used in the manufacture of medicaments for
CC	the treatment or prevention of infection or disease caused by
CC	Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC	Nucleic acids encoding (I) are used to detect Streptococcus in a
CC	biological sample. (I) is used to determine whether a compound binds to
CC	(I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC	used as a vaccine or diagnostic composition. The disease caused by
CC	Streptococcus that is prevented or treated may be meningitis. Nucleic
CC	acid encoding (I) may be used to recombinantly produce (I) and may be
CC	used in gene therapy. Antibodies to (I) are used for affinity
CC	chromatography, immunoassays, and distinguishing/identifying
CC	Streptococcus proteins
XX	
SO	Sequence 939 BP; 257 A; 172 C; 201 G; 309 T; 0 U; 0 Other;
	Query Match 49.4%; Score 23.2; DB 6; Length 939;
	Best Local Similarity 70.5%; Pred. No. 2.8e+02;
	Matches 31; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Oy	3 ACAGGCTGATTAGAAATAAATTGATGACACATTTCGAATTT 46
Db	791 ACAAAGCATGAACAAAGACAGTTTGATTACATTTCACATGT 834
RESULT 20	
ACD13402	
ID	ACD13402 standard; cDNA, 1978 BP.
XX	
AC	ACD13402;
XX	
DT	13-AUG-2003 (first entry)
XX	
DB	Human DNA encoding a p53 modifier. SEQ ID 74.
XX	
KW	Human; ss; gene; p53 modifier; cytosolic; cancer; cytosolic;
KW	antiangiogenic; antiapoptotic; p53 pathway; breast cancer; colon cancer;
KW	kidney cancer; lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW	apoptotic disorder; cell proliferation disorder.
XX	
OS	Homo sapiens.
XX	
PN	MO200299122-A1.
XX	
DD	12-DEC-2002.
XX	
PF	03-JUN-2002; 2002WO-US017382.

XX	05-JUN-2001; 2001US-0296076P.
PR	10-OCT-2001; 2001US-0328605P.
FR	15-FEB-2002; 2002US-0357253P.
XX	
PA	(EXEL-) EXELIXIS INC.
PI	Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
DR	WPI; 2003-156859/15.
XX	P-PSDB; ABO07229.
PT	Identifying modulators of the p53 pathway for use in treating apoptotic
PT	or cell proliferation disorders, comprises screening for agents that
PT	modulate activity of a human ortholog of genes that modify the p53
PT	pathway in Drosophila.
XX	
PS	Example 2; Page 357-358; 678pp; English.
XX	
CC	The invention relates to identifying (M1) a candidate p53 pathway
CC	modulating agent, by contacting an assay system comprising a purified HM
CC	polypeptide (human ortholog of genes that modify the p53 pathway in
CC	Drosophila) or nucleic acid with a test agent under conditions, where but
CC	for the presence of the test agent, the system provides a reference
CC	activity, and detecting a test agent-biased activity of the assay system.
CC	Also included are modulating (M2) a p53 pathway of a cell (comprising
CC	contacting a cell defective in p53 function with a candidate modulator
CC	that specifically binds to a HM polypeptide comprising an HM amino acid
CC	sequence, where p53 function is restored), modulating (M3) a p53 pathway
CC	in a mammalian cell (comprising contacting the cell with an agent that
CC	specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
CC	a disease in a patient (comprising: (a) obtaining a biological sample
CC	from the patient; (b) contacting the sample with a probe for HM
CC	expression; (c) comparing the results with a control; and (d) determining
CC	whether the comparison indicates a likelihood disease). (M1) is useful
CC	for identifying modulators of the p53 pathway. A probe for HM expression
CC	is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
CC	in a patient, where the cancer has greater than 25 % expression level.
CC	Modulators identified by (M1) are useful in a variety of diagnostic and
CC	therapeutic applications, where disease or disorder prognosis is related
CC	to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
CC	proliferation disorders (e.g. cancer). Another two new methods (M2 and
CC	M3) are useful for modulating the p53 pathway of a cell, thus restoring
CC	the p53 function of the cell, so that the cell undergoes normal
CC	proliferation or progression through the cell cycle. (M2) and (M3) are
CC	also useful for treating defects in the p53 pathway such as angiogenic,
CC	apoptotic or cell proliferation disorders. The present sequence is an HM
CC	nucleic acid encoding a p53 pathway modifying protein
XX	
SQ	Sequence 1978 BP; 527 A; 476 C; 556 G; 419 T; 0 U; 0 Other;
	Query Match            49.4%; Score 23.2; DB 8; Length 1978;
	Best Local Similarity    83.3%; Pred. No. 3e+02;
	Matches     25; Conservative    1; Mismatches    4; Indels       0; Gaps       0;
Oy	10 TGATTAGAAATTAASITTTGATCACCATTTT 39                                   Db     1939 TGACTTGAAATAAATTGGAACAACATTTT 1968
RESULT 21	
ID	ACN41054
XX	ACN41054 standard; cDNA; 1978 BP.
XX	
AC	ACN41054;
XX	
DT	18-NOV-2004 (first entry)
XX	
DE	Tumour-associated antigenic target (TAT) cDNA DNA270496, SEQ ID NO:6239.
KX	Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KX	tumour; diagnosis; cell proliferative disorder; breast cancer;
KX	colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW	



KW central nervous system cancer; bladder cancer; pancreatic cancer;  
KW cervical cancer; melanoma; leukaemia; hybridisation probe;  
KW chromosome identification; chromosome mapping; gene mapping;  
KW gene therapy; cytostatic; gene; ss.  
XX Homo sapiens.  
XX MO2004030615-A2.  
XX 15-APR-2004.  
XX 29-SEP-2003; 2003WO-US028547.  
XX 02-OCT-2002; 2002US-0414971P.  
XX (GETH ) GENENTECH INC.  
XX Wu TD, Zhang Z, Zhou Y;  
XX MPI: 2004-347921/32.  
XX P-PSDB; ABM82429.  
XX New tumor-associated antigenic target polypeptides and nucleic acids,  
XX useful in preparing a medicament for treating or detecting a  
XX proliferative disorder, e.g. breast, lung, colorectal, ovarian or  
XX prostate cancer or tumor.  
XX Claim 1; SEQ ID NO 6239; 7273bp; English.  
XX  
XX The invention relates to human tumor-associated antigenic target (TAT)  
XX polypeptides, and their related nucleic acids. The TAT polypeptides are  
XX overexpressed in cancer tissues compared to normal tissues, and may thus  
XX serve as effective targets for the diagnosis and treatment of cancer in  
XX mammals. The invention also relates to nucleic acid and polypeptide  
XX sequences at least 80% identical to the TAT nucleic acids and  
XX polypeptides; expression vectors and host cells comprising a TAT nucleic  
XX acid; an antibody specific for a TAT polypeptide; a peptide or organic  
XX molecule which binds to a TAT polypeptide; fusion proteins comprising a  
XX TAT polypeptide; and methods and compositions for the treatment or  
XX diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,  
XX antibodies, antagonists, binding molecules and compositions are useful  
XX for diagnosing or treating a cell proliferative disorder associated with  
XX increased TAT expression, particularly cancers such as breast cancer,  
XX colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder  
XX cancer, pancreatic cancer, cervical cancer, cancers of the central  
XX nervous system, melanoma and leukaemia. TAT nucleic acids may further be  
XX used as hybridisation probes, in chromosome and gene mapping, in  
XX chromosome identification and in gene therapy. The present sequence  
XX represents a TAT nucleic acid of the invention  
XX  
XX Sequence 1978 BP; 527 A; 476 C; 556 G; 419 T; 0 U; 0 Other;  
SQ  
XX  
XX Query Match 49.4%; Score 23.2; DB 13; Length 1978;  
XX Best Local Similarity 83.3%; Pred. No. 3e+02;  
XX Matches 25; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 10 TGATTGAAATTAAGTTGATCACCATTTT 39  
DB 1939 TGACTTGAATTAATTAACCTTGACACCAATTTT 1968  
RESULT 22  
ADJ40809  
XX ADJ40809 standard; cDNA; 2000 BP.  
XX  
XX ADJ40809;  
AC  
XX 06-MAY-2004 (first entry)  
XX  
XX Plant CDNA #1809.  
DE  
XX  
XX Plant; gene; ss; transcriptions; plant genome augmentation; cereal;  
KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;

KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;  
KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;  
KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;  
KW antifungal.  
XX Eukaryota.  
XX  
XX US2004016025-A1.  
XX 22-JAN-2004.  
XX 26-SEP-2002; 2002US-00260238.  
XX 26-SEP-2001; 2001US-0325277P.  
XX 26-SEP-2001; 2001US-0325448P.  
XX 04-APR-2002; 2002US-0370620P.  
XX (BUDW/) BUDWORTH P.  
XX (MOUG/) MOUGHAMER T.  
XX (BRIG/) BRIGGS S P.  
XX (COOP/) COOPER B.  
XX (GLAZ/) GLAZEBROOK J.  
XX (GOFF/) GOFF S A.  
XX (KATG/) KATAGIRI F.  
XX (KREP/) KREPS J.  
XX (PROV/) PROVART N.  
XX (RICK/) RICHE D.  
XX (ZHU/) ZHU T.  
XX Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J,  
XX Goff SA, Katagiri F, Krep J, Provart N, Riche D, Zhu T;  
XX MPI: 2004-190374/18.  
XX  
XX New rice promoter, useful for manipulating crop plants to alter or  
XX improve phenotypic characteristics, e.g. produce large quantities of oil  
XX or proteins, resistance to insecticides, virus or fungi, stress tolerance  
XX or high nutritional value.  
XX  
XX Claim 26; SEQ ID NO 1809; 230bp; English.  
XX  
XX The invention relates to plant nucleotide sequences that direct seed-,  
XX leaf- and/or stem-, panicle-, root- or pollen-specific or preferential  
XX or constitutive transcription of an operatively linked nucleic acid  
XX segment. The invention also relates to a method for augmenting a plant  
XX genome and a method of identifying a gene, where its expression is  
XX altered in the seed, leaf, stem, panicle, pollen, root or is constitutive  
XX in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,  
XX canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,  
XX sorghum, rice or wheat. The polynucleotides and the polypeptides they  
XX encode are useful for manipulating crop plants to alter or improve  
XX phenotypic characteristics, to produce large quantities of oil or  
XX proteins, to incur resistance to insecticides, viruses or fungi, and to  
XX incur stress tolerance (e.g. salt, cold or drought) to ensure the plants  
XX have a high nutritional value with reduced apical dominance or dwarfism,  
XX early flowering or altered metabolic pathways. This sequence represents a  
XX plant nucleic acid of the invention. Note: The sequence data for this  
XX patent did not form part of the printed specification but was obtained in  
XX electronic format directly from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX  
XX Sequence 2000 BP; 614 A; 364 C; 405 G; 617 T; 0 U; 0 Other;  
SQ  
XX  
XX Query Match 49.4%; Score 23.2; DB 12; Length 2000;  
XX Best Local Similarity 77.8%; Pred. No. 3e+02;  
XX Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 12 ATTGAAATTAAGTTGATCACCATTTCGAATTTT 47  
DB 788 ACTTAATATCATTAAACACCATTTCGAATTTT 823  
RESULT 23  
ADM02167

```

XX ADW02167 standard; cDNA; 2347 BP.
XX AC ADM02167;
XX DT 20-MAY-2004 (first entry)
XX DE Human cDNA of the invention SEQ ID NO:852.
XX KW ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX OS Homo sapiens.
XX EN EP1347046-A1.
XX PD 24-SEP-2003.
XX PF 12-APR-2002; 2002EP-00008400.
XX PR 22-MAR-2002; 2002JP-00137785.
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX PI Isogai T, Sugiyama T, Otsuki T, Makamatsu A, Sato H, Ishii S,
XX PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
XX PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y,
XX DR WPI; 2003-723558/69.
XX DR P-PSDB; ADM04610.
XX PT New polynucleotides and polypeptides are useful in gene therapy, for
XX PT developing a diagnostic marker or medicines for regulating their
XX PT expression and activity, or as a target of gene therapy.
XX PS Claim 1; SEQ ID NO 852; 3055pp; English.
XX CC The invention relates to a novel human polynucleotide and the encoded
XX CC polypeptide. A polynucleotide of the invention ADM06202-ADM06773 is useful
XX CC as a primer for synthesizing the polynucleotide or as a probe for
XX CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
XX CC useful in gene therapy, for developing a diagnostic marker or medicines
XX CC for regulating their expression and activity, or as a target of gene
XX CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
XX CC are useful as pharmaceutical agents. The present sequence represents a
XX CC cDNA sequence of the invention.
XX SQ Sequence 2347 BP; 583 A; 516 C; 580 G; 668 T; 0 U; 0 Other;

Query Match 49.4%; Score 23.2; DB 11; Length 2347;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 25; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 10 TGATTGAAATTAATTGATCACCATTTT 39
DB 2308 TGACTGAAATTAACCTTGAAACACAAATTT 2337

RESULT 24
ABL28314
ID ABL28314 standard; DNA; 3212 BP.
XX AC ABL28314;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 36415.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; de.
XX OS Drosophila melanogaster.
XX EN WO200171042-A2.
```

```

XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 36415; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 3212 BP; 849 A; 631 C; 681 G; 1051 T; 0 U; 0 Other;

Query Match 49.4%; Score 23.2; DB 4; Length 3212;
Best Local Similarity 70.5%; Pred. No. 3.2e+02;
Matches 31; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 4 CAGCGTTGATTGAAATTAATTGATCACCATTTCATTAATTT 47
DB 2108 CACATTGATTAAAAATGATTTTGTGCTACCTAATTTAATTTT 2151

RESULT 25
AAK89062/C
ID AAK89062 standard; DNA; 5468 BP.
XX AC AAK89062;
XX DT 05-NOV-2001 (first entry)
XX DE Human digestive system antigen genomic sequence SEQ ID NO: 2638.
XX KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
XX KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
XX KW digestive system disorder; Meckel's diverticulum; ds.
XX OS Homo sapiens.
XX EN WO200155314-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001324.
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205515P.
XX PR 07-JUN-2000; 2000US-0209467P.
```

PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226799P.  
PR 22-AUG-2000; 2000US-0226811P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 08-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234977P.  
PR 25-SEP-2000; 2000US-0234988P.  
PR 26-SEP-2000; 2000US-0235848P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241212P.  
PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249255P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0251989P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JUN-2001; 2001US-0259678P.  
  
(HUMA-) HUMAN GENOME SCI INC.  
XX PA  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-502630/55.  
XX  
XX  
PT Polynucleotides encoding digestive system antigens, useful for  
XX diagnosing, treating, preventing and/or prognosing disorders of the  
XX digestive system, particularly cancer and cancer metastases.  
PS Disclosure; SEQ ID NO 2638; 986bp; English.  
XX  
XX  
CC The present invention provides the protein and coding sequences of a  
CC number of human digestive system antigens. These can be used in the  
CC diagnosis, treatment and prevention of digestive system disorders,  
CC including cancer, Meckel's diverticulum, bacterial or parasitic  
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or  
CC ulcerative colitis. The present sequence is a genomic DNA fragment

CC encoding a digestive system antigen of the invention

XX Sequence 5468 BP; 1838 A; 981 C; 1052 G; 1597 T; 0 U; 0 Other;

Query Match 49.4%; Score 23.2; DB 4; Length 5468;

Best Local Similarity 83.3%; Pred. No. 3.4e+02; Mismatches 4; Indels 0; Gaps 0;

Oy 10 TGATTAGAAATAAASTTTGATCACCATTTT 39

Db 1914 TGACTTGAATAAATACTTGACACCAATTTT 1885

Search completed: June 2, 2005, 08:57:27

Job time : 466 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2005, 08:20:08 / Search time 3128 Seconds  
(without alignments)  
571.938 Million cell updates/sec

Title: US-10-349-143-1

Perfect score: 47

Sequence: 1 ccacagcgtcgtatagaat.....gattcaccatttcaatttt 47

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

EST:\*  
1: gb\_esc1:\*  
2: gb\_esc2:\*  
3: gb\_esc3:\*  
4: gb\_esc4:\*  
5: gb\_esc5:\*  
6: gb\_esc6:\*  
7: gb\_esc7:\*  
8: gb\_esc8:\*  
9: gb\_esc9:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27.8	59.1	463	8	AQ076343 CIT-HSP-2
2	27.6	58.7	945	9	CL473755 SAIL_205
3	26.4	56.2	1227	9	AG333342 Mus muscu
4	26.2	55.7	456	9	BX906921 Leishmani
5	26.2	55.7	489	8	BZ196382 CH230-322
6	26	55.3	633	2	BS524921 BS524921
7	26	55.3	2375	3	AK086348 Mus muscu
8	26	55.3	3058	3	AK043320 Mus muscu
9	25.8	54.9	349	1	AJ503513 AJ503513
10	25.8	54.9	507	1	AJ503688 AJ503688
11	25.6	54.5	765	9	CR071307 Reverse b
12	25.6	54.5	416	4	BI402346 MT-P-CPO-
13	25.6	54.5	422	7	CV294559 EST882936
14	25.6	54.5	443	7	CV300373 EST880889
15	25.6	54.5	453	7	CV294479 EST882856
16	25.6	54.5	462	7	CV295114 EST883491
17	25.6	54.5	493	7	CV293479 EST881856
18	25.6	54.5	495	7	CV300458 EST880974
19	25.6	54.5	507	7	CV300453 EST880969
20	25.6	54.5	540	7	CV295268 EST883645
21	25.6	54.5	555	7	CV294106 EST882483
22	25.6	54.5	614	7	CV297810 EST886187
23	25.6	54.5	620	7	CV293638 EST882015
24	25.6	54.5	630	7	CV294470 EST882847

25	25.6	54.5	638	7	CV293898	CV293898 EST882275
26	25.6	54.5	650	7	CV295618	CV295618 EST883995
27	25.6	54.5	658	7	CV297609	CV297609 EST885985
28	25.6	54.5	671	7	CV295246	CV295246 EST883623
29	25.6	54.5	673	7	CV298905	CV298905 EST887364
30	25.6	54.5	676	7	CV295713	CV295713 EST884090
31	25.6	54.5	755	5	BM971114	BM971114 UI-CF-EC1
32	25.4	54.0	758	5	BM417776	BM417776 BM417776
33	25.4	54.0	1182	8	CC297513	CC297513 CH261-105
34	25.2	53.6	548	5	BP052517	BP052517 BP052517
35	25.2	53.6	609	1	AL655320	AL655320 AL655320
36	25.2	53.6	836	8	BZ571268	BZ571268 m82_1809
37	25	53.2	607	8	BZ169970	BZ169970 CH230-414
38	25	53.2	1101	9	CNS609FJ	AL071159 Drosophila
39	24.8	52.8	272	2	BR338231	BR338231 BR338231
40	24.8	52.8	488	8	B41058	B41058 HS-1052-B2-
41	24.8	52.8	521	5	BP181191	BP181191 BP181191
42	24.8	52.8	601	7	CNS33313	CNS33313 UI-M-HO-
43	24.8	52.8	639	2	BB252497	BB252497 BB252497
44	24.8	52.8	773	8	BH498017	BH498017 BQJX32TF
45	24.8	52.8	707	8	BH551295	AZ533670 ENTBT50TR
46	24.8	52.8	746	8	CH118714	CL021677 CH216-8L2
47	24.8	52.8	756	8	BH577556	AK043061 Mus muscu
48	24.8	52.8	773	8	BH498017	BY600070 BY600070
49	24.8	52.8	916	8	AZ533670	AQ370013 HS-5045-A
50	24.8	52.8	1072	3	CL021677	BQ487419 WHE1713-1
51	24.8	52.8	3367	9	AK043061	AZ157787 SP 0058 A
52	24.6	52.3	470	6	BY600070	BX180649 Dario r6T
53	24.6	52.3	498	5	AQ370013	CNS03796 EC2BBA8CH
54	24.6	52.3	504	5	BQ487419	CG988894 CH240_145
55	24.6	52.3	519	4	BM557140	AA768788 OC17909_8
56	24.6	52.3	635	6	CB507009	BM557140 BM557140
57	24.6	52.3	831	9	BX180649	CB507009 bsalob505
58	24.4	51.9	475	1	AA768788	BO273840 rc71d01.Y
59	24.4	51.9	493	5	BM557140	BZ939750 CH240_106
60	24.4	51.9	635	6	CB507009	CNS09138 EC2BBA31A
61	24.4	51.9	662	5	BO273840	CA063687 s6a1rg5b3
62	24.4	51.9	702	7	BZ939750	AZ263847 RPT1-23-1
63	24.4	51.9	725	7	CNS09138	BY691928 BY691928
64	24.4	51.9	747	6	CA063687	CR068163 Reverse b
65	24.4	51.9	778	6	AZ263847	BX954363 DXFZP781P
66	24.2	51.5	198	8	AZ263847	AZ262790 RPT1-23-1
67	24.2	51.5	405	6	BY691928	CL592069 OB_Ba005
68	24.2	51.5	443	9	CR068163	BH199785 Sm1-39A11
69	24.2	51.5	450	5	BX954363	CE205000 t1gr-g88-
70	24.2	51.5	474	6	AZ262790	BU433553 603219945
71	24.2	51.5	556	-9	CL592069	AZ870628 ZK0183K12
72	24.2	51.5	587	8	BH199785	CL133935 ISB1-104D
73	24.2	51.5	589	9	CE205000	BU849838 AGENCOURT
74	24.2	51.5	657	5	BU433553	CC085229 GCU-K33r.
75	24.2	51.5	708	8	AZ870628	CL649974 CH213-264
76	24.2	51.5	879	9	CL133935	CD507583 CNA86-A06
77	24.2	51.5	946	5	BU849838	CR749668 Homo. Bap1
78	24.2	51.5	972	8	CC085229	CV293480 EST881857
79	24.2	51.5	1122	6	CL649974	BP334038 BP334038
80	24.2	51.5	1170	3	CR507583	BH037779 RPT1-24-3
81	24.2	51.5	1576	6	CR749668	AO384426 RPT11-13
82	24	51.1	483	7	CV293480	CB634378 OSI1EB13N
83	24	51.1	583	5	BP334038	CC004237 PUEB014TD
84	24	51.1	621	8	BH037779	CC205996 CH261-56U
85	24	51.1	647	8	AO384426	D18433 MUCS01491
86	24	51.1	734	6	CB634378	CO728416 UGC-Bend
87	24	51.1	734	6	CB634378	DI8433 MUCS01491
88	24	51.1	734	6	CB634378	CR854930 SWMB33CAW
89	23.8	50.6	360	7	CK150865	CK150865 CmpgEST02
90	23.8	50.6	377	7	CK150865	BY615187 BY615187
91	23.8	50.6	420	7	CK150865	CL349441 RPT14-30
92	23.8	50.6	431	6	BY615187	CK703544 ZP101-P00
93	23.8	50.6	471	9	CL349441	BM571363 BM571363
94	23.8	50.6	471	9	CL349441	BH760753 BMBAC316P
95	23.8	50.6	500	5	BM571363	
96	23.8	50.6	513	8	BH760753	
97	23.8	50.6	513	8	BH760753	

	23.8	50.6	533	8	A0524290	HS 5237	B
98	23.8	50.6	533	8	A0524290	HS 5237	B
99	23.8	50.6	548	4	B0692271	B0692271	
100	23.8	50.6	554	4	B0531372	B0531372	PIC1_65_C

## ALIGNMENTS

RESULT 1	A0076343	463 bp	DNA	linear	GSS, 20-AUG-1998
LOCUS	A0076343				
DEFINITION	CIT-HSP-2559G24.TF CIT-HSP Homo sapiens genomic clone 2559G24,				

ACCESSION	A0076343	
VERSION	A0076343.1	GI:3437527
KEYWORDS	GSS.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eultheria; Primates; Carnivora; Hominoidea; Homo.	
REFERENCE	1 (bases 1 to 463)	
AUTHORS	Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.	

TITLE	Use of a random human BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL	Unpublished (1998)
COMMENT	Other_GSSs: CIT-HSP-2359G24.TR

Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel.: 301 838 0200  
Fax: 301 838 0208  
Email: [mdadam@tigr.org](mailto:mdadam@tigr.org)  
Clones are available from Research Genetics ([info@resgen.com](mailto:info@resgen.com)). BAC  
end search page:  
[http://www.tigr.org/cdb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html).  
Seq primer: M13-21  
Class: BAC ends.

```

FEATURES
    source
        location/Qualifiers
            1..463
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /clone="2359624"
                /sex="Male"
                /cell_type="Sperm"
                /clone_1fb="Citr-HSP"
                /note="Vector: pBeloBAC11, site_1: HindIII, site_2: HindIII"

```

Query Match	59.1%	Score 27.8	DB 8	Length 463
Best Local Similarity	78.0%	Pred. No. 1,2e+02		
Matches	32	Conservative	1	Mismatches 8
				Indels 0
				Gaps 0
QY	6	GGCTTGATTCAGATAAATTTGATCACCATTTCGAAATTT	46	
Db	421	GGCTTGCTTCAGATAAATCTGATTAACATTTTAAATTT	461	

RESULT 2	
CL473755/c	
LOCUS	CL473755
DEFINITION	945 bp DNA linear GSS 01-APR-2004
LOCUS	SAIL_205_E07.v1 SAIL Collection Arabidopsis thaliana genomic clone
DEFINITION	SAIL_205_E07.v1, genomic survey sequence.
LOCUS	Arabidopsis thaliana

KEYWORDS  
GSS.  
SOURCE  
Arabidopsis thaliana (thale cress)  
ORGANISM  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 945)

REFERENCE

Sessions, A., Burke, E., Preeing, G., Aux, G., McElver, J., Patton, D.,

AUTHORS

TITLE	A high-throughput Arabidopsis reverse genetics system
JOURNAL	Plant Cell 14 (12), 2985-2994 (2002)
MEDLINE	22356987
PUBMED	12468722
COMMENT	Contact: Sessions A,

Syngenta Biotechnology Inc  
3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA  
Email: allen.sessions@syngenta.com  
ABRC Stock Number CS809613; T-DNA left border flanking sequences of  
Syngenta Arabidopsis insertion library (SAIL) lines are available  
through the Arabidopsis Biological Resource Center (ABRC).  
Sequences represent a pool of amplified genomic regions and not  
single contiguous sequences.  
Class: TDNA tagged.

FEATURES	source
Location/Qualifiers	1..945
/organism="Arabidopsis thaliana"	
/mol_type="genomic DNA"	
/ecotype="Columbia"	
/db_xref="taxon:3702"	
/clone="SAIL_205_E07.v1"	
/clone_1lb="SAIL Collection"	
/note="1-DNA left border sequences were isolated using a modified Tail-PCR strategy"	

```
Query Match      58.7%; Score 27.6; DB 9; Length 945;  
Best Local Similarity   75.0%; Pred. No. 1.4e+02;  
Matches    33; Conservative     1; Mismatches    10; Indels       0;  
  
QY          4 CAGCGTTGATTACGAATTAASSTTGGTCACCATTCTTCAAAATTT 47  
            ||||| :|||||  
Db         375 CAGCGTTGTAGAACAATAAGACTTGATCCTTGATTAGAACATTT 332
```

RESULT 3	AG333342	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
	AG333342	1227 bp	DNA	linear	GSS 02--JUN-2004			
	Mus musculus	molossinus	DNA, clone:MS901-123N09.TJ, genomic survey	sequence.	AG333342			
	AG333342.1	GI:47906652						
	GSS.							
	Mus musculus	molossinus						
	Mus musculus	molossinus						

REFERENCE	1
AUTHORS	Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
TITLE	BAC end Sequences of Library MSNg01
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1227)
AUTHORS	Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical

COMMENT  
Clones are derived from the mouse BAC library Msmgvi. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp), Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan  
phone: 81-298-36-9189, fax: 81-298-36-9199

e-mail: abe@rc.riken.jp  
PRIMERS  
Sequencing : TJ  
LIBRARY : PBACe3.6  
Vector : ECoRI  
R.Site 1 : ECoRI  
R.Site 2 : ECoRI  
Location/Qualifiers  
1..1227  
/organism="Mus musculus molossinus"  
/mol\_type="genomic DNA"  
/sub\_species="molossinus"  
/db\_xref="taxon:57486"  
/clone="MSMG01-123N09.TJ"  
/sex="male"  
/tissue\_type="mixture of kidney and spleen"  
/clone\_1ib="MSMG01 Mouse Male BAC library"

ORIGIN  
Query Match 56.2%; Score 26.4; DB 9; Length 1227;  
Best Local Similarity 71.7%; Pred.No.3.5e+02;  
Matches 33; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

Oy 2 CACAGCGTTGATTAGAAATPAAATTTGATCACCATTTCGAATTTT 47  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 98 CACAAGCTTACTTAGACTACCTCTGTGTACCATCTTGAATTTT 143

RESULT 4  
BX906921/c 456 bp DNA linear GSS 07-JAN-2004  
LOCUS Leishmania braziliensis GSS, clone LBAF74F03, genomic survey  
DEFINITION  
sequence.  
ACCESSION BX906921 GI:40735389  
VERSION GSS: genomic survey sequence.  
KEYWORDS Leishmania braziliensis  
SOURCE Leishmania braziliensis  
ORGANISM Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;  
Leishmania; Leishmania braziliensis species complex.  
REFERENCE  
1 Laurentino,E.C., Ruiz,J.C. and Cruz,A.K.  
AUTHORS GSS analysis of the Leishmania braziliensis genome  
TITLE Unpublished  
JOURNAL 2 (bases 1 to 456)  
REFERENCE Cruz,A.K.  
AUTHORS Direct Submission  
TITLE Submitted (07-JAN-2004) Cruz A.K., University of Sao Paulo,  
JOURNAL Department of Molecular and Cell Biology, FMRP, Avenida  
Bandeirantes 3900, Ribeirao Preto, SP, 14049-900, BRAZIL  
COMMENT Clone requests: akcruz@fmrp.usp.br.  
FEATURES  
Location/Qualifiers  
1..456  
/organism="Leishmania braziliensis"  
/mol\_type="genomic DNA"  
/strain="WHOM/BR/75/W2904"  
/db\_xref="taxon:5660"  
/clone="LBAF74F03"

ORIGIN  
Query Match 55.7%; Score 26.2; DB 9; Length 456;  
Best Local Similarity 79.5%; Pred.No.3.9e+02;  
Matches 31; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 1 CCACAGCGTTGATTAGAAATPAAATTTGATCACCATTTT 39  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 127 CCACAGCGATGAAAAAAGTAAATTTGACCCCATTTT 89

RESULT 5  
BZ196382 489 bp DNA linear GSS 11-OCT-2002  
LOCUS CH230-322A14.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone  
DEFINITION

CH230-322A14, genomic survey sequence.  
BZ196382  
BZ196382.1 GI:23854434  
GSS.  
Rattus norvegicus (Norway rat)  
KEYWORDS Rattus norvegicus  
SOURCE Rattus norvegicus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 489)  
Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,  
AUTHORS Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,  
Riggs,F., de Jong,P. and Fraser,C.M.  
TITLE Rat BAC End Sequences from library CHORI-230 MboI segment  
JOURNAL Unpublished (1999)  
COMMENT Other\_GSSs: CH230-322A14.TJ  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@ligr.org  
Clones are derived from the rat BAC library CHORI-230  
(http://www.chori.org/bacpac/rat230.htm). For BAC library  
availability, please contact Pieter de Jong (pdejong@mail.cho.org).  
Clones may be purchased from BACPAC Resources  
(http://www.chori.org/bacpac/or ering\_information.htm). BAC end  
page: http://www.ligr.org/cdb/bac\_ends/rat/bac\_end\_intro.html  
plate: 322 row: A column: 14  
Seq primer: SP6  
Class: BAC ends.  
FEATURES  
Location/Qualifiers  
1..489  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/strain="BN/SNHsd/MCW"  
/db\_xref="taxon:10116"  
/clone="CH230-322A14"  
/sex="female"  
/cell\_type="Brain"  
/clone\_1ib="CHORI-230 Segment 2"  
/note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;  
CHORI-230 Rat (BN/SNHsd/MCW) BAC library produced by  
Pieter de Jong"

ORIGIN  
Query Match 55.7%; Score 26.2; DB 8; Length 489;  
Best Local Similarity 75.6%; Pred.No.3.9e+02;  
Matches 31; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Oy 7 GCTTGATTGAAATPAAATTTGATCACCATTTCGAATTTT 47  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 385 GCATGGTTGATCCATCTTGTATCACCATTTCATATCTT 425

RESULT 6  
BB524921 633 bp mRNA linear EST 25-OCT-2001  
LOCUS BB524921 RIKEN full-length enriched, 15 days embryo head Mus  
DEFINITION musculus cDNA clone D930023E13 3', mRNA sequence.  
ACCESSION BB524921  
VERSION BB524921.2 GI:16444057  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 633)  
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,  
REFERENCE Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,  
AUTHORS Kono,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,  
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,





TITLE	COMMENT	FEATURES	ORIGIN	Qy	Db	RESULT 8	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
Muramatsu, M. and Hayashizaki, Y.	Submitted (16-ARR-2002) Yoshitake Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.	1. 2375 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:D930023E13" /db_xref="taxon:10090" /clone="D930023E13" /tissue_type="head" /clone_id="RIKEN full-length enriched mouse cDNA library" /dev_stage="15 days embryo" 688..987 /note="unnamed protein product; 5-HYDROXYTRYPTAMINE 2C RECEPTOR (SWISSPROT P34968, evidence: FASTA, 99.6%ID, 100%length, match=1377) putative" /codon_start=1 /protein_id="BAC39653.1" /db_xref="GI:26352033" /translation="MVNIGTRVRSLSVHLIGLWQEPDISPSPAAYTTPFNSDGG RLFPDPGVQMPALSTVIVIIIMTIGNIIIVIMVSMKKLHAINYFLMSLGHC"	12 ATTGAAATTAATTTGATCACCATTTCATTTTCAATTTT 47       	2231 ATTGCAATTAACTTGAACCTTGATTTCATTTT 2196	AK043320/c	AK043320	Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone: A730084N12 product: 5-HYDROXYTRYPTAMINE 2C RECEPTOR, full insert sequence.	AK043320	AK043320.1 GI:26089632	HTC; CAP trapper.	Mus musculus (house mouse)	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)	1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	

JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	
AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitanishi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kasai, K., Fujiwara, S., Inoue, K., Togawa, Y., Itawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuzaki, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	
AUTHORS	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
TITLE	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
JOURNAL	6 (Pages 1 to 3058)
REFERENCE	
AUTHORS	Adachi, J., Aizawa, K., Akiyama, T., Arikawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagiri, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirotsane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyu, S., Kurihara, C., Matsumura, T., Miyazaki, A., Murata, M., Nakamura, Y., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohashi, N., Ohashi, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tanagami, M., Tagawa, A., Takahashi, F., Takai-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp). URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers
FEATURES	1. 3058 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="PANTOM_DB:A730084N12" /db_xref="taxon:10090" /clone="A730084N12" /clone_type="cerebellum" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="7 days neonate" 1. 3058 /note="5-HYDROXYTRYPTAMINE 2C RECEPTOR (SWISSPROT P34968, evidence: FASTV, 99.6%ID, 100%length, match-1377)"
ORIGIN	
Query Match	55.3%; Score 26; DB 3; Length 3058;

Best Local Similarity 80.6%; Pred. No. 4.9e+02;  
Matches 29; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```

QY      12 ATTGAATAAATTTGATCACCATTTCGAATTTT 47
          ||||| : ||||| ||||| |||||
Db      2229 ATTAGCAATTAACCTTGAACTGATTTTCAAAATTTT 2194

```

RESULT 9	LOCUS	DEFINITION	LOCUS	DEFINITION
AJ503513	AJ503513	349 bp	truncatula	CDNA clone mtgmadc120035a04,
				mRNA sequence.

ACCESSION	AJ503513
VERSION	AJ503513.1
KEYWORDS	GI:22084441
SOURCE	EST.
ORGANISM	Medicago truncatula (barrel medic)
	Medicago truncatula

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.

REFERENCE	AUTHORS	TITLE
1 (Pages 1 to 349)	Manthey,K., Bartelmeier,V., Baier,M.C., Meyer,F., Bartels,D., Babel,T., Linke,B., Grunwald,U., Franken,P., Kueser,H., Perlick,A.M. and Puhler,A.	Detection of transript sequences from mycorrhizal roots of the

**JOURNAL COMMENT**  
model mycorrhiza Medicago truncatula genotype Al7 - Glomus intraradices using the approach of an EST genome project  
Unpublished (2002)  
Contact: Kuester H

LEHRSTUHL FÜR GENETIK  
UNIVERSITÄT BIELEFELD  
POSTFACH 100131, D-33501 BIELEFELD, GERMANY.  
LOCATION/QUALIFIERS  
1. .349

```

/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/cisone="mtcmadcl20035a04"
/tissue_type="mycorrhizal roots"
/dev_stage="six week old mycorrhizal roots harvested five
weeks after inoculation with Glomus intraradices"
/clone_id="MTAMP"
/note="Vector: pGEM-T; Site 1: PstI; Site 2: SphI;
genotype A17; cDNA was prepared from polyA+ enriched RNA
from mycorrhizal roots harvested five weeks after
inoculation. The cDNA was directionally ligated by
Medigenomix into the pGEM-T vector from Promega using
GATCGCGCCGAGCGCGCCACATG and CTCGAGCCCATATGCGCCGCG
adapters. Plasmids containing cDNA inserts were propagated
in E. coli DH10B cells."

```

ORIGIN			
Query Match	54.9%	Score 25.8;	DB 1; Length 349;
Best Local Similarity	76.9%;	Pred. No. 5.1e+02;	
Matches 30;	Conservative 1;	Mismatches 0;	Gaps 0;

Oy 5 AGGCTTGATTAGAAATATAAATTGGTTCACCATTTTCAA 43  
||| ||| | : | ||| ||| |||  
Db 71 AGGGTTCATAAAAATTCACCCTTATCACCCATTTTCAAA 109

RESULT 10					
AJ503688					
LOCUS					
AJ503688					
507 bp					
mRNA					
linear					
EST 10-FEB-2003					

DEFINITION	AJ503688 MTMP Medicago truncatula cDNA clone mtgmacd120037a12, mRNA sequence.
ACCESSION	AJ503688
VERSION	AJ503688.1
KEYWORDS	GI:22084616
	EST_3

SOURCE	Medicago truncatula (barrel medic)
ORGANISM	Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE  
AUTHORS

Medicago.  
1 (bases 1 to 507)  
Manthey, K., Bartelsmeier, V., Baier, M. C., Meyer, F., Bartels, D.

TITLE	Detection of transcript sequences from mycorrhizal roots of the model mycorrhiza Medicago truncatula genotype A17 - Glomus intraradices using the approach of an EST genome project
JOURNAL	Unpublished (2002)
COMMENT	Contact: Kuester H

**FEATURES**  
source location/Qualifiers  
1. .507

```
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="caxon:3880"
/clone="mtgmadc120037a12"
```

```

/tissue_type="mycorrhizal roots"
/dev_stage="six week old mycorrhizal roots harvested five
weeks after inoculation with Glomus intraradices"
/clone_id="MTAMP"
/note="Vector: pGEM-T; Site 1: PstI; Site 2: SphI;

```

genotype A17: cDNA was prepared from polyA<sup>+</sup> enriched RNA from mycorrhizal roots harvested five weeks after inoculation. The cDNA was directionally ligated by MedGenomix into the pGEM-T vector from Promega using

ORIGIN  
GATACCGGCCGAGCGCGCCGACATG and CTGACGCCCATTATGCGCGG  
GATACCGGCCGAGCGCGCCGACATG and CTGACGCCCATTATGCGCGG  
adapters. Plasmids containing cDNA inserts were propagated  
in *E. coli* DH10B cells."

	54.9% ; Score 25.8; DB 1; Length 507;
Query Match	Best Local Similarity 76.9%; Pred.No.5.2e+02;
Matches	30; Conservative 1; Mismatches 8; Indels 0; Gaps 0
Oy	5 AGGCTTGATTGAATAAATTTGCACCACTTTCAAA 43                                          Db   71 AAGGTTCATAAAAAAATTCACCCTTAATCACACTTTCCAAA 109
RESULT 11	
CRO7I1307/c	
LOCUS	CRO7I1307           765 bp       DNA       linear   GSS 05-JUL-2000
DEFINITION	Reverse strand read from insert in 5'HPT insertion targeting and chromosome engineering clone MHPN20114, genomic survey sequence.
ACCESSION	CRO7I1307
VERSION	CRO7I1307.1 GI:49804897

KEYWORDS	SSS; genome survey sequence; MISC.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Euteleostomi; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE	AUTHORS
1 (bases 1 to 765)	
	Adams, D. J., Biggs, P. J., Cox, A. V., Davies, R. M., van der Weyden, L., Jonkers, J., Smith, J., Plumb, R. W., Taylor, R. G., Nishijima, I., Yu, Y., Rogers, J., and Bradley, A.

TITLE	Direct Submission
JOURNAL	Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. <a href="http://www.sanger.ac.uk/MICR">http://www.sanger.ac.uk/MICR</a>
FEATURES	Location/Qualifiers
source	1..765

```
/organism="Mus musculus"  
/mol_type="genomic DNA"  
/db_xref="taxon:10090"  
/clone_1b="MHEN201114"  
/clone_1b="MHEN"
```

ORIGIN

Query Match 54.9%; Score 25.8; DB 9; Length 765;  
 Best Local Similarity 76.9%; Pred. No. 5.3e+02;  
 Matches 30; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 5 AGGCTTGATTAGAAATTAAGTTGATCACCATTTCAAA 43  
 DB 716 AAGCTTGAAATAGAGATTAAGTTGGAGACTATTTCAGA 678

RESULT 12 BI402346 416 bp mRNA linear EST 14-AUG-2001  
 LOCUS BI402346/c MI-P-CPD-nw-a-03-0-UI.61 MI-P-CPD Sus scrofa cDNA clone  
 DEFINITION MI-P-CPD-nw-a-03-0-UI 3', mRNA sequence.  
 ACCESSION BI402346 GI:15181407  
 VERSION BI402346.1  
 KEYWORDS EST.  
 SOURCE Sus scrofa (pig)  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 416)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8889548

COMMENT Contact: Tugale CK  
 Molecular Genetics Laboratory, Department of Animal Science  
 Iowa State University  
 201 Kildee Hall, Ames, IA 50011-3150, USA  
 Tel: 5152944252  
 Fax: 5152942401  
 Email: Cktugale@iastate.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the non-normalized uterus library cDNA library Preparation: M.B. Soares Lab, University of Iowa EST sequencing: M.B. Soares Lab, University of Iowa Clone distribution: clones will be available through Research Genetics (www.reagen.com) The following repetitive elements were found in this cDNA sequence: 1-24,  
 >AT rich#low complexity  
 Seq primer: M13 Forward  
 POLYA=Yes.

#### FEATURES

source location/Qualifiers  
 1..416  
 /organism="Sus scrofa"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9823"  
 /clone="MI-P-CPD-nw-a-03-0-UI"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="MI-P-CPD"  
 /note="Vector: pTR73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: EcoRI; The MI-P-CPD library is derived from uteri. For a detailed description of the library from which this clone was derived, please visit our web site at <http://piggen.genome.iastate.edu/>. The procedure used to create this library has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)  
 TAG\_TISSUE=uterus  
 TAG\_LIB=MI-P-CPD  
 TAG\_SEQ=AGTCCATCG"

#### ORIGIN

Query Match 54.5%; Score 25.6; DB 4; Length 416;  
 Best Local Similarity 82.4%; Pred. No. 6e+02;  
 Matches 28; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 10 TGATTAGAAATTAAGTTGATCACCATTTCAAA 43  
 DB 48 TGACTGAAATTAAGCTTGACACCAATTTTAAAA 15

RESULT 13 CV294559 422 bp mRNA linear EST 23-SEP-2004  
 LOCUS CV294559/c EST882936 petunia floral post-ethylene cDNA library Petunia x hybrida cDNA clone Petunia-C2H4-20-G12 5' end, mRNA sequence.  
 DEFINITION hybrida cDNA clone Petunia-C2H4-20-G12 5' end, mRNA sequence.  
 ACCESSION CV294559  
 VERSION CV294559.1 GI:52583969  
 KEYWORDS EST.  
 SOURCE Petunia x hybrida  
 ORGANISM Petunia x hybrida  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Petunia.

REFERENCE 1 (bases 1 to 422)  
 AUTHORS Shibusaki,K., Underwood,B., Loucas,H., Farnerie,W., Jones,M. and Clark,D.  
 TITLE Petunia x hybrida EST collection  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: David Clark  
 UF Floriculture Biotechnology Lab  
 University of Florida  
 Environmental Horticulture Department, 1545 Fifield Hall, Box 110670, Gainesville, FL 32611-0670, USA  
 Tel: 352-392-1831 x370  
 Fax: 352-392-3870  
 Email: dclark@mail.ufl.edu  
 Contact Dr. Clark (dclark@mail.ufl.edu) for clone information  
 Seq primer: T3 primer.

#### FEATURES

source location/Qualifiers  
 1..422  
 /organism="Petunia x hybrida"  
 /mol\_type="mRNA"  
 /cultivar="Mitchell Diploid (aka. Mitchell, aka W115 in Europe)"  
 /db\_xref="taxon:4102"  
 /clone="petunia-C2H4-20-G12"  
 /tissue\_type="all floral organs"  
 /lab\_host="lambda ZAPI unidirectional"  
 /clone\_lib="petunia floral post-ethylene cDNA library"  
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; supplier: Petunia x hybrida cv. Mitchell Diploid plants were grown from seeds to a fully flowering stage under standard greenhouse conditions. Flowers at anthesis stage were excised from plants grown in standard greenhouses and treated with 2ppm exogenous ethylene. Entire flowers were collected at 0, 5, 10, 15, 24, and 36 hours after treatment. Total RNA was extracted from each sample, and 100 micrograms of each sample was combined for subsequent poly A+ mRNA selection and cDNA synthesis."

#### ORIGIN

Query Match 54.5%; Score 25.6; DB 7; Length 422;  
 Best Local Similarity 73.8%; Pred. No. 6e+02;  
 Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 6 GGCTTGATTAGAAATTAAGTTGATCACCATTTCAAATTTT 47  
 DB 327 GGCTGATCAACAATAATGATTGATCACCCTTCACAAATGTT 286

RESULT 14 CV300373 443 bp mRNA linear EST 23-SEP-2004  
 LOCUS CV300373/c EST880889 petunia floral development cDNA library Petunia x hybrida  
 DEFINITION EST880889 petunia floral development cDNA library Petunia x hybrida  
 ACCESSION CV300373  
 VERSION CV300373.1 GI:52595614  
 KEYWORDS EST.

Query Match 54.5%; Score 25.6; DB 7; Length 422;  
 Best Local Similarity 73.8%; Pred. No. 6e+02;  
 Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

SOURCE  
ORGANISM Petunia x hybrida  
Petunia x hybrida  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Petunia.  
REFERENCE  
AUTHORS 1 (bases 1 to 443)  
Shibuya,K., Underwood,B., Loucas,H., Farmerie,W., Jones,M. and  
Clark,D.  
TITLE Petunia x hybrida EST collection  
JOURNAL Unpublished (2004)  
COMMENT Contact: David Clark  
UP floriculture Biotechnology Lab  
University of Florida  
Environmental Horticulture Department, 1545 Field Hall, Box  
110670, Gainesville, FL 32611-0670, USA  
Tel: 352-392-1831 x370  
Fax: 352-392-3870  
Email: dclark@mail.ifas.ufl.edu  
Contact Dr. Clark (dclark@mail.ifas.ufl.edu) for clone information  
Seq primer: T3 primer.  
Location/Qualifiers  
FEATURES  
source  
1..443  
/organism="Petunia x hybrida"  
/mol\_type="mRNA"  
/cultivar="Mitchell Diploid (aka. Mitchell, aka W15 in Europe)"  
/db\_xref="taxon:4102"  
/clone="Petunia-resql-D04"  
/tissue\_type="all floral organs"  
/lab\_host="lambda ZAPit unidirectional"  
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; supplier: Petunia x hybrida cv. Mitchell Diploid plants were grown from seeds to a fully flowering stage under standard greenhouse conditions. Ten entire flowers of six developmental stages were collected on the same day from plants grown in standard greenhouses. The flower stages were as follows in chronological order from youngest to oldest: stage 1 - no color in corolla; corolla 0.5 inches long stage 2 - first sign of color in corolla; corolla .75-1 inches long stage 3 - fully elongated corolla (not open); corolla 1.5 inches long stage 4 - fully open corolla; anthers not yet dehiscent stage 5 - fully open corolla; freshly anthesed, bright yellow pollen; wet stigma stage 6 - pre-senescent; yellowing of corolla tube; dry brown pollen (if present); stigma dry. Total RNA was extracted from each sample, and 100 micrograms of each sample was combined for subsequent poly A+ mRNA selection and cDNA synthesis."

ORIGIN  
Query Match 54.5%; Score 25.6; DB 7; Length 443;  
Best Local Similarity 73.8%; Pred. No. 66+02; Mismatches 10; Indels 0; Gaps 0;  
Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
QY 6 GGCTTGATTAGAAATTAATTGATCACCATTTCGAATTTT 47  
DB 257 GGCTGCATCACAATAATTGATTGATCACCCTTCACAAATGTT 216

RESULT 15  
CV294479/c 453 bp mRNA linear EST 23-SEP-2004  
LOCUS ESR982856 petunia floral post-ethylene cDNA library Petunia x  
DEFINITION hybrida cDNA clone Petunia-C2H4-20-A01 5' end, mRNA sequence.  
ACCESSION CV294479  
VERSION CV294479.1 GI:52583811  
KEYWORDS EST.  
SOURCE Petunia x hybrida  
ORGANISM Petunia x hybrida  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Petunia.

REFERENCE  
AUTHORS 1 (bases 1 to 453)  
Shibuya,K., Underwood,B., Loucas,H., Farmerie,W., Jones,M. and  
Clark,D.  
TITLE Petunia x hybrida EST collection  
JOURNAL Unpublished (2004)  
COMMENT Contact: David Clark  
UP floriculture Biotechnology Lab  
University of Florida  
Environmental Horticulture Department, 1545 Field Hall, Box  
110670, Gainesville, FL 32611-0670, USA  
Tel: 352-392-1831 x370  
Fax: 352-392-3870  
Email: dclark@mail.ifas.ufl.edu  
Contact Dr. Clark (dclark@mail.ifas.ufl.edu) for clone information  
Seq primer: T3 primer.  
Location/Qualifiers  
FEATURES  
source  
1..453  
/organism="Petunia x hybrida"  
/mol\_type="mRNA"  
/cultivar="Mitchell Diploid (aka. Mitchell, aka W15 in Europe)"  
/db\_xref="taxon:4102"  
/clone="Petunia-C2H4-20-A01"  
/tissue\_type="all floral organs"  
/lab\_host="lambda ZAPit unidirectional"  
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; supplier: Petunia x hybrida cv. Mitchell Diploid plants were grown from seeds to a fully flowering stage under standard greenhouse conditions. Flowers at anthesis stage were excised from plants grown in standard greenhouses and treated with 2ppm exogenous ethylene. Entire flowers were collected at 0, 5, 10, 15, 24, and 36 hours after treatment. Total RNA was extracted from each sample, and 100 micrograms of each sample was combined for subsequent poly A+ mRNA selection and cDNA synthesis."

ORIGIN  
Query Match 54.5%; Score 25.6; DB 7; Length 453;  
Best Local Similarity 73.8%; Pred. No. 66+02; Mismatches 10; Indels 0; Gaps 0;  
Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
QY 6 GGCTTGATTAGAAATTAATTGATCACCATTTCGAATTTT 47  
DB 327 GGCTGCATCACAATAATTGATTGATCACCCTTCACAAATGTT 286

RESULT 16  
CV295114 462 bp mRNA linear EST 23-SEP-2004  
LOCUS ESR983491 petunia floral post-ethylene cDNA library Petunia x  
DEFINITION hybrida cDNA clone Petunia-C2H4-25-A04 5' end, mRNA sequence.  
ACCESSION CV295114  
VERSION CV295114.1 GI:52585086  
KEYWORDS EST.  
SOURCE Petunia x hybrida  
ORGANISM Petunia x hybrida  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamiales; Solanales; Solanaceae; Petunia.  
REFERENCE  
AUTHORS 1 (bases 1 to 462)  
Shibuya,K., Underwood,B., Loucas,H., Farmerie,W., Jones,M. and  
Clark,D.  
TITLE Petunia x hybrida EST collection  
JOURNAL Unpublished (2004)  
COMMENT Contact: David Clark  
UP floriculture Biotechnology Lab  
University of Florida  
Environmental Horticulture Department, 1545 Field Hall, Box  
110670, Gainesville, FL 32611-0670, USA  
Tel: 352-392-1831 x370  
Fax: 352-392-3870  
Email: dclark@mail.ifas.ufl.edu





LOCUS CV294106 555 bp mRNA linear EST 23-SEP-2004  
DEFINITION EST862483 petunia floral post-ethylene cDNA library Petunia x  
hybrida cDNA clone Petunia-C2H4-16-G01 5' end, mRNA sequence.  
ACCESSION CV294106  
VERSION CV294106.1 GI:52583065  
KEYWORDS EST.  
SOURCE Petunia x hybrida  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Petunia.  
1 (bases 1 to 555)  
REFERENCE Shibusya,K., Underwood,B., Loucas,H., Farmerie,W., Jones,M. and  
Clark,D.  
AUTHORS  
TITLE Petunia x hybrida EST collection  
JOURNAL Unpublished (2004)  
COMMENT Contact: David Clark  
UF Horticulture Biotechnology Lab  
University of Florida  
Environmental Horticulture Department, 1545 Fifield Hall, Box  
110670, Gainesville, FL 32611-0670, USA  
Tel: 352-392-1831 x370  
Fax: 352-392-3870  
Email: dclark@mail.ifas.ufl.edu  
Contact Dr. Clark (dclark@mail.ifas.ufl.edu) for clone information  
Seg primer: T3 primer.  
Location/Qualifiers  
1. .555  
/organism="Petunia x hybrida"  
/mol\_type="mRNA"  
/cultivar="Mitchell diploid (aka. Mitchell, aka W15 in  
Europe)"  
/db\_xref="taxon:4102"  
/clone="Petunia-C2H4-16-G01"  
/tissue\_type="all floral organs"  
/lab\_host="lambda ZAPII unidirectional"  
/clone\_lib="Petunia floral post-ethylene cDNA library"  
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:  
XhoI; supplier: Petunia x hybrida cv. Mitchell diploid  
plants were grown from seeds to a fully flowering stage  
under standard greenhouse conditions. Flowers at anthesis  
stage were excised from plants grown in standard  
greenhouses and treated with 2ppm exogenous ethylene.  
Entire flowers were collected at 0, 5, 10, 15, 24, and 36  
hours after treatment. Total RNA was extracted from each  
sample, and 100 micrograms of each sample was combined for  
subsequent poly A+ mRNA selection and cDNA synthesis."

ORIGIN  
Query Match 54.5%; Score 25.6; DB 7; Length 555;  
Best Local Similarity 73.8%; Pred. No. 6.1e+02;  
Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 6 GGCTGATTAGAAATTAATTTGATCACCATTTCAAATTTT 47  
||||| ||||| :||||||| ||||| |||||  
Db 264 GGCTGCATCACAAAATTAATTGATTCACCTTCACAAATGTT 223

RESULT 22  
LOCUS CV297810 614 bp mRNA linear EST 23-SEP-2004  
DEFINITION EST866187 petunia floral development cDNA library Petunia x hybrida  
cDNA clone Petunia-DeVA-23-A10 5' end, mRNA sequence.  
ACCESSION CV297810  
VERSION CV297810.1 GI:52590467  
KEYWORDS EST.  
SOURCE Petunia x hybrida  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Petunia.  
1 (bases 1 to 614)  
REFERENCE Shibusya,K., Underwood,B., Loucas,H., Farmerie,W., Jones,M. and  
AUTHORS

TITLE Petunia x hybrida EST collection  
JOURNAL Unpublished (2004)  
COMMENT Contact: David Clark  
UF Horticulture Biotechnology Lab  
University of Florida  
Environmental Horticulture Department, 1545 Fifield Hall, Box  
110670, Gainesville, FL 32611-0670, USA  
Tel: 352-392-1831 x370  
Fax: 352-392-3870  
Email: dclark@mail.ifas.ufl.edu  
Contact Dr. Clark (dclark@mail.ifas.ufl.edu) for clone information  
Seg primer: T3 primer.  
Location/Qualifiers  
1. 614  
/organism="Petunia x hybrida"  
/mol\_type="mRNA"  
/cultivar="Mitchell diploid (aka. Mitchell, aka W15 in  
Europe)"  
/db\_xref="taxon:4102"  
/clone="Petunia-DeVA-23-A10"  
/tissue\_type="all floral organs"  
/lab\_host="lambda ZAPII unidirectional"  
/clone\_lib="Petunia floral development cDNA library"  
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:  
XhoI; supplier: Petunia x hybrida cv. Mitchell diploid  
plants were grown from seeds to a fully flowering stage  
under standard greenhouse conditions. Ten entire flowers  
of six developmental stages were collected on the same day  
from plants grown in standard greenhouses. The flower  
stages were as follows in chronological order from  
youngest to oldest: stage 1 - no color in corolla; corolla  
0.5 inches long stage 2 - first sign of color in corolla;  
corolla .75-1 inches long stage 3 - fully elongated  
corolla (not open); corolla 1.5 inches long stage 4 -  
fully open corolla; anthers not yet dehiscent stage 5 -  
fully open corolla; freshly anthesis, bright yellow  
pollen; wet stigma stage 6 - pre-anthesis; yellowing of  
corolla tube; dry brown pollen (if present); stigma dry.  
Total RNA was extracted from each sample, and 100  
micrograms of each sample was combined for subsequent poly  
A+ mRNA selection and cDNA synthesis."

ORIGIN  
Query Match 54.5%; Score 25.6; DB 7; Length 614;  
Best Local Similarity 73.8%; Pred. No. 6.1e+02;  
Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 6 GGCTGATTAGAAATTAATTTGATCACCATTTCAAATTTT 47  
||||| ||||| :||||||| ||||| |||||  
Db 250 GGCTGCATCACAAAATTAATTGATTCACCTTCACAAATGTT 209

RESULT 23  
LOCUS CV293638 620 bp mRNA linear EST 23-SEP-2004  
DEFINITION EST82015 petunia floral post-ethylene cDNA library Petunia x  
hybrida cDNA clone Petunia-C2H4-8RR-D11 5' end, mRNA sequence.  
ACCESSION CV293638  
VERSION CV293638.1 GI:52582127  
KEYWORDS EST.  
SOURCE Petunia x hybrida  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Petunia.  
1 (bases 1 to 620)  
REFERENCE Shibusya,K., Underwood,B., Loucas,H., Farmerie,W., Jones,M. and  
Clark,D.  
AUTHORS  
TITLE Petunia x hybrida EST collection  
JOURNAL Unpublished (2004)  
COMMENT Contact: David Clark  
UF Horticulture Biotechnology Lab



University of Florida  
Environmental Horticulture Department, 1545 Fifield Hall, Box  
110670, Gainesville, FL 32611-0670, USA  
Tel: 352-392-1831 x370  
Fax: 352-392-3870  
Email: dclark@mail.ufl.edu  
Contact Dr. Clark (dclark@mail.ufl.edu) for clone information  
Seq primer: T3 primer.

## FEATURES

## source

1..620  
/organism="Petunia x hybrida"  
/mol\_type="mRNA"  
/cultivar="Mitchell Diploid (aka. Mitchell, aka W15 in Europe)"  
/db\_xref="taxon:4102"  
/clone="Petunia-C2H4-8R-D11"  
/tissue\_type="all floral organs"  
/lab\_host="lambda ZAP11 unidirectional"  
/clone\_lib="petunia floral post-ethylene cDNA library"  
/note="vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; supplier: Petunia x hybrida cv. Mitchell Diploid plants were grown from seeds to a fully flowering stage under standard greenhouse conditions. Flowers at anthesis stage were excised from plants grown in standard greenhouses and treated with 2ppm exogenous ethylene. Entire flowers were collected at 0, 5, 10, 15, 24, and 36 hours after treatment. Total RNA was extracted from each sample, and 100 micrograms of each sample was combined for subsequent poly A+ mRNA selection and cDNA synthesis."

## ORIGIN

Query Match 54.5%; Score 25.6; DB 7; Length 620;  
Best Local Similarity 73.8%; Pred. No. 6.1e+02;  
Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

6 GCCTTGATTGAAATTAATTTGATCACCATTTCGAATTTT 47  
243 GGCTGCATCACAAATTAAGTTGATCACCCTTCACAAATGTT 202

RESULT 24  
CV294470/c 630 bp mRNA linear EST 23-SEP-2004  
LOCUS  
DEFINITION EST882847 petunia floral post-ethylene cDNA library Petunia x hybrida cDNA clone Petunia-C2H4-21-H04 5' end, mRNA sequence.  
ACCESSION CV294470  
VERSION CV294470.1 GI:52583793  
KEYWORDS  
SOURCE  
ORGANISM  
Petunia x hybrida  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Petunia.  
1 (bases 1 to 630)  
Shibuya,K., Underwood,B., Loucas,H., Farmerie,W., Jones,M. and Clark,D.  
Petunia x hybrida EST collection  
Unpublished (2004)  
Contact: David Clark  
UF Floriculture Biotechnology Lab  
University of Florida  
Environmental Horticulture Department, 1545 Fifield Hall, Box 110670, Gainesville, FL 32611-0670, USA  
Tel: 352-392-1831 x370  
Fax: 352-392-3870  
Email: dclark@mail.ufl.edu  
Contact Dr. Clark (dclark@mail.ufl.edu) for clone information  
Seq primer: T3 primer.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Unpublished (2004)  
Contact: David Clark  
UF Floriculture Biotechnology Lab  
University of Florida  
Environmental Horticulture Department, 1545 Fifield Hall, Box 110670, Gainesville, FL 32611-0670, USA  
Tel: 352-392-1831 x370  
Fax: 352-392-3870  
Email: dclark@mail.ufl.edu  
Contact Dr. Clark (dclark@mail.ufl.edu) for clone information  
Seq primer: T3 primer.

## FEATURES

## source

1..630  
/organism="Petunia x hybrida"  
/mol\_type="mRNA"  
/cultivar="Mitchell Diploid (aka. Mitchell, aka W15 in Europe)"

Europe)"  
/db\_xref="taxon:4102"  
/clone="Petunia-C2H4-21-H04"  
/tissue\_type="all floral organs"  
/lab\_host="lambda ZAP11 unidirectional"  
/clone\_lib="petunia floral post-ethylene cDNA library"  
/note="vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; supplier: Petunia x hybrida cv. Mitchell Diploid plants were grown from seeds to a fully flowering stage under standard greenhouse conditions. Flowers at anthesis stage were excised from plants grown in standard greenhouses and treated with 2ppm exogenous ethylene. Entire flowers were collected at 0, 5, 10, 15, 24, and 36 hours after treatment. Total RNA was extracted from each sample, and 100 micrograms of each sample was combined for subsequent poly A+ mRNA selection and cDNA synthesis."

## ORIGIN

Query Match 54.5%; Score 25.6; DB 7; Length 630;  
Best Local Similarity 73.8%; Pred. No. 6.1e+02;  
Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

6 GCCTTGATTGAAATTAATTTGATCACCATTTCGAATTTT 47  
327 GGCTGCATCACAAATTAAGTTGATCACCCTTCACAAATGTT 286

RESULT 25  
CV293898/c 638 bp mRNA linear EST 23-SEP-2004  
LOCUS  
DEFINITION EST882275 petunia floral post-ethylene cDNA library Petunia x hybrida cDNA clone Petunia-C2H4-15-D09 5' end, mRNA sequence.  
ACCESSION CV293898  
VERSION CV293898.1 GI:52582649  
KEYWORDS  
SOURCE  
ORGANISM  
Petunia x hybrida  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Petunia.  
1 (bases 1 to 638)  
Shibuya,K., Underwood,B., Loucas,H., Farmerie,W., Jones,M. and Clark,D.  
Petunia x hybrida EST collection  
Unpublished (2004)  
Contact: David Clark  
UF Floriculture Biotechnology Lab  
University of Florida  
Environmental Horticulture Department, 1545 Fifield Hall, Box 110670, Gainesville, FL 32611-0670, USA  
Tel: 352-392-1831 x370  
Fax: 352-392-3870  
Email: dclark@mail.ufl.edu  
Contact Dr. Clark (dclark@mail.ufl.edu) for clone information  
Seq primer: T3 primer.

## FEATURES

## source

1..638  
/organism="Petunia x hybrida"  
/mol\_type="mRNA"  
/cultivar="Mitchell Diploid (aka. Mitchell, aka W15 in Europe)"  
/db\_xref="taxon:4102"  
/clone="Petunia-C2H4-15-D09"  
/tissue\_type="all floral organs"  
/lab\_host="lambda ZAP11 unidirectional"  
/clone\_lib="petunia floral post-ethylene cDNA library"  
/note="vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; supplier: Petunia x hybrida cv. Mitchell Diploid plants were grown from seeds to a fully flowering stage under standard greenhouse conditions. Flowers at anthesis stage were excised from plants grown in standard greenhouses and treated with 2ppm exogenous ethylene. Entire flowers were collected at 0, 5, 10, 15, 24, and 36



hours after treatment. Total RNA was extracted from each sample, and 100 micrograms of each sample was combined for subsequent poly A+ mRNA selection and cDNA synthesis."

ORIGIN

Query Match 54.5%; Score 25.6; DB 7; Length 638;  
 Best Local Similarity 73.8%; Pred. No. 6.1e+02;  
 Matches 31; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
 QY 6 GGCTTGATTAGAAATTAATTAATTTGATCACCATTTCGAATTT 47  
 ||||| ||| | ||| : ||||| || ||||| ||  
 Db 629 GGCTCGATCACAATAATTAATTTGATCACCCTTCACAATGTT 588

Search completed: June 2, 2005, 09:49:49  
 Job time : 3145 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 2, 2005, 08:30:33 ; Search time 126 Seconds  
(without alignments)  
610.357 Million cell updates/sec

Title: US-10-349-143-1

Perfect score: 47

Sequence: 1 ccacagcttgattagaat.....gatcacatttcaatttt 47

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/6C\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/6D\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46.6	99.1	47	US-09-422-978-1	Sequence 1, Appl1
2	25.2	53.6	601	US-09-949-016-52779	Sequence 52779, A
3	25.2	53.6	601	US-09-949-016-52780	Sequence 52780, A
4	25.2	53.6	187580	US-09-949-016-13266	Sequence 13266, A
5	24.6	52.3	37195	US-09-949-016-13264	Sequence 13264, A
6	24.2	51.5	273	US-09-428-796A-7359	Sequence 7359, A
7	24.2	51.5	993	US-09-328-352-3876	Sequence 3876, A
8	24	51.1	96922	US-09-949-016-17061	Sequence 17061, A
9	23.6	50.2	601	US-09-949-016-52894	Sequence 52894, A
10	23.6	50.2	90876	US-09-949-016-13271	Sequence 13271, A
11	23.2	49.4	601	US-09-949-016-185469	Sequence 185469, A
12	23.2	49.4	601	US-09-949-016-185470	Sequence 185470, A
13	23.2	49.4	1978	US-09-949-016-1358	Sequence 1358, A
14	23.2	49.4	2064	US-09-248-796A-1897	Sequence 1897, A
15	23.2	49.4	24538	US-09-949-016-13100	Sequence 13100, A
16	23.2	49.4	44971	US-09-949-016-17049	Sequence 17049, A
17	23	48.9	76164	US-09-949-016-12288	Sequence 12288, A
18	23	48.9	76165	US-09-949-016-14005	Sequence 14005, A
19	22.8	48.5	31491	US-09-360-186-1	Sequence 1, Appl1
20	22.8	48.5	31491	US-09-864-680A-1	Sequence 1, Appl1
21	22.6	48.1	459	US-08-956-171E-1138	Sequence 1138, A
22	22.6	48.1	459	US-08-781-986A-1138	Sequence 1138, A
23	22.6	48.1	601	US-09-949-016-14510	Sequence 14510, A
24	22.6	48.1	1209	US-08-336-165A-55	Sequence 55, Appl1
25	22.6	48.1	57726	US-09-949-016-16430	Sequence 16430, A
26	22.6	48.1	61178	US-09-949-016-17369	Sequence 17369, A
27	22.6	48.1	181429	US-09-949-016-12372	Sequence 12372, A

28	22.6	48.1	181430	US-09-949-016-15772	Sequence 15772, A
29	22.6	48.1	205163	US-09-949-016-17009	Sequence 17009, A
30	22.6	48.1	236964	US-09-949-016-15753	Sequence 15753, A
31	22.4	47.7	348	US-09-583-110-1030	Sequence 1030, A
32	22.4	47.7	492	US-09-107-433-1285	Sequence 1285, A
33	22.4	47.7	745	US-09-495-050A-183	Sequence 183, A
34	22.4	47.7	3257	US-09-949-016-1496	Sequence 1496, A
35	22.4	47.7	3986	US-09-949-016-24	Sequence 24, Appl1
36	22.4	47.7	46603	US-09-949-016-15636	Sequence 15636, A
37	22.4	47.7	63558	US-09-949-016-13238	Sequence 13238, A
38	22.4	47.7	64489	US-09-949-016-11766	Sequence 11766, A
39	22.2	47.2	601	US-09-949-016-30145	Sequence 30145, A
40	22.2	47.2	601	US-09-949-016-37382	Sequence 37382, A
41	22.2	47.2	601	US-09-949-016-65350	Sequence 65350, A
42	22.2	47.2	601	US-09-949-016-142406	Sequence 142406, A
43	22.2	47.2	601	US-09-949-016-14595	Sequence 14595, A
44	22.2	47.2	981	US-09-671-317-38	Sequence 38, Appl1
45	22.2	47.2	3001	US-09-539-333D-127	Sequence 127, Appl1
46	22.2	47.2	28213	US-09-949-016-12738	Sequence 12738, A
47	22.2	47.2	28216	US-09-949-016-13652	Sequence 13652, A
48	22.2	47.2	55886	US-09-949-016-15129	Sequence 15129, A
49	22.2	47.2	62311	US-09-949-016-14582	Sequence 14582, A
50	22.2	47.2	95890	US-09-949-016-16412	Sequence 16412, A
51	22.2	47.2	211049	US-09-949-016-15770	Sequence 15770, A
52	22	46.8	571	US-09-854-133-307	Sequence 307, Appl1
53	22	46.8	601	US-09-949-016-124854	Sequence 124854, A
54	22	46.8	601	US-09-949-016-136288	Sequence 136288, A
55	22	46.8	601	US-09-949-016-151005	Sequence 151005, A
56	22	46.8	49416	US-09-949-016-15234	Sequence 15234, A
57	22	46.8	50950	US-09-949-016-16659	Sequence 16659, A
58	22	46.8	91772	US-09-949-016-15568	Sequence 15568, A
59	22	46.8	132438	US-09-949-016-14349	Sequence 14349, A
60	22	46.8	132438	US-09-949-016-14350	Sequence 14350, A
61	22	46.8	151809	US-09-949-016-14348	Sequence 14348, A
62	22	46.8	192302	US-09-949-016-15270	Sequence 15270, A
63	22	46.8	239527	US-09-949-016-15980	Sequence 15980, A
64	22	46.8	524032	US-09-949-016-16928	Sequence 16928, A
65	22	46.8	524032	US-09-949-016-16929	Sequence 16929, A
66	22	46.8	524032	US-09-949-016-16930	Sequence 16930, A
67	22	46.8	524032	US-09-949-016-16931	Sequence 16931, A
68	22	46.8	529885	US-09-949-016-14340	Sequence 14340, A
69	22	46.8	529885	US-09-949-016-14341	Sequence 14341, A
70	22	46.8	529885	US-09-949-016-14342	Sequence 14342, A
71	22	46.8	529885	US-09-949-016-14343	Sequence 14343, A
72	22	46.8	529885	US-09-949-016-14344	Sequence 14344, A
73	22	46.8	529885	US-09-949-016-14345	Sequence 14345, A
74	22	46.8	529885	US-09-949-016-14346	Sequence 14346, A
75	22	46.8	529885	US-09-949-016-14347	Sequence 14347, A
76	21.8	46.4	472	US-08-956-171E-948	Sequence 948, Appl1
77	21.8	46.4	472	US-08-949-001-140	Sequence 140, Appl1
78	21.8	46.4	601	US-09-949-001-140	Sequence 140, Appl1
79	21.8	46.4	601	US-09-949-001-140	Sequence 140, Appl1
80	21.8	46.4	861	US-09-276-531-90	Sequence 90, Appl1
81	21.8	46.4	1368	US-09-540-236-731	Sequence 731, Appl1
82	21.8	46.4	5208	US-08-781-891-77	Sequence 77, Appl1
83	21.8	46.4	5208	US-09-791-211-11	Sequence 11, Appl1
84	21.8	46.4	5208	US-09-618-166-70	Sequence 70, Appl1
85	21.8	46.4	5208	US-09-949-001-9	Sequence 9, Appl1
86	21.8	46.4	5208	US-09-949-001-9	Sequence 9, Appl1
87	21.8	46.4	5208	US-09-949-001-9	Sequence 9, Appl1
88	21.8	46.4	5208	US-09-949-001-9	Sequence 9, Appl1
89	21.8	46.4	5208	US-09-949-001-9	Sequence 9, Appl1
90	21.8	46.4	5208	US-09-949-001-9	Sequence 9, Appl1
91	21.8	46.4	5208	US-09-949-001-9	Sequence 9, Appl1
92	21.8	46.4	5208	US-09-949-001-9	Sequence 9, Appl1
93	21.8	46.4	5208	US-09-949-001-9	Sequence 9, Appl1
94	21.8	46.4	5208	US-09-949-001-9	Sequence 9, Appl1
95	21.8	46.4	5208	US-09-949-001-9	Sequence 9, Appl1
96	21.8	46.4	5208	US-09-949-001-9	Sequence 9, Appl1
97	21.8	46.4	5208	US-09-949-001-9	Sequence 9, Appl1
98	21.8	46.4	5208	US-09-949-001-9	Sequence 9, Appl1
99	21.8	46.4	5208	US-09-949-001-9	Sequence 9, Appl1
100	21.8	46.4	5208	US-09-949-001-9	Sequence 9, Appl1

## ALIGNMENTS

RESULT 1  
US-09-422-978-1; Sequence 1, Application US/09422978  
; Patent No. 6537751

; GENERAL INFORMATION:

; APPLICANT: Cohen, Daniel

; APPLICANT: Blumenfeld, Marta

; APPLICANT: Chumakov, Ilya

; TITLE OF INVENTION: Biallelic markers for use in constructing a high density....

; FILE REFERENCE: GENSET.020CPI

; CURRENT APPLICATION NUMBER: US/09/422,978

; PRIOR FILING DATE: 1999-10-20

; EARLIER APPLICATION NUMBER: US 09/298,850

; EARLIER FILING DATE: 1999-04-21

; EARLIER APPLICATION NUMBER: US 60/109,732

; EARLIER FILING DATE: 1998-11-23

; EARLIER APPLICATION NUMBER: US 60/082,614

; NUMBER OF SEQ ID NOS: 11796

; SEQ ID NO 1

; LENGTH: 47

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE:

; NAME/KEY: allele

; LOCATION: 24

; OTHER INFORMATION: 99-109-224 : polymorphic base G or C

US-09-422-978-1

Query Match

Best Local Similarity 99.1%; Score 46.6; DB 4; Length 47;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match

Best Local Similarity 100.0%; Pred. No. 2.4e-07;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match

Best Local Similarity 100.0%; Pred. No. 2.4e-07;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match

Best Local Similarity 100.0%; Pred. No. 2.4e-07;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match

Best Local Similarity 100.0%; Pred. No. 2.4e-07;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match

Best Local Similarity 100.0%; Pred. No. 2.4e-07;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match

Best Local Similarity 100.0%; Pred. No. 2.4e-07;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match

Best Local Similarity 100.0%; Pred. No. 2.4e-07;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match

Best Local Similarity 100.0%; Pred. No. 2.4e-07;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match

Best Local Similarity 100.0%; Pred. No. 2.4e-07;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 296 CACCACTGTTAGAAATATAAGTTCATGCCCATTTCAAAATTT 341

RESULT 3  
US-09-949-016-52780; Sequence 52780, Application US/09949016  
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 52780

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-52780

Query Match

Best Local Similarity 53.6%; Score 25.2; DB 4; Length 601;

Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Query Match

Best Local Similarity 71.7%; Pred. No. 11;

Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Query Match

Best Local Similarity 71.7%; Pred. No. 24;

Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Query Match

Best Local Similarity 71.7%; Pred. No. 24;

Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Query Match

Best Local Similarity 71.7%; Pred. No. 24;

Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Query Match

Best Local Similarity 71.7%; Pred. No. 24;

Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Query Match

Best Local Similarity 71.7%; Pred. No. 24;

Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Query Match

Best Local Similarity 71.7%; Pred. No. 24;

Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Query Match

Best Local Similarity 71.7%; Pred. No. 24;

Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Query Match

Best Local Similarity 71.7%; Pred. No. 24;

Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Query Match

Best Local Similarity 71.7%; Pred. No. 24;

Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Query Match

Best Local Similarity 71.7%; Pred. No. 24;

Matches 33; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Db 178619 CACCATCATGTTTAAATAATATAGTTCAGCCCATTTTCAATTCT 178574

## RESULT 5

US-09-949-016-13264/c  
; Sequence 13264, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13264  
; LENGTH: 37195  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(37195)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13264

Query Match 52.3%; Score 24.6; DB 4; Length 37195;  
Best Local Similarity 70.2%; Pred. No. 31;  
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CCACAGGCTTGATAGAAATAASTTTGATCACCATTTCGAATTTT 47  
Db 10945 CCAGAGGCTTGATGCTGGGAAAGAAATGATACCATTTTGAATTTT 10899

## RESULT 6

US-09-248-796A-7359/c  
; Sequence 7359, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinslock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; PRIOR FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 7359  
; LENGTH: 273  
; TYPE: DNA  
; ORGANISM: Candida albicans  
US-09-248-796A-7359

Query Match 51.5%; Score 24.2; DB 4; Length 273;  
Best Local Similarity 68.1%; Pred. No. 21;  
Matches 32; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CCACAGGCTTGATAGAAATAASTTTGATCACCATTTCGAATTTT 47  
Db 211 CCACAGTTTATTGGAAATAATGCGATCCCAACTTCAATCTT 165

## RESULT 7

US-09-328-352-3876/c  
; Sequence 3876, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; PRIOR FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 3876  
; LENGTH: 993  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-3876

Query Match 51.5%; Score 24.2; DB 4; Length 993;  
Best Local Similarity 68.1%; Pred. No. 25;  
Matches 32; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CCACAGGCTTGATAGAAATAASTTTGATCACCATTTCGAATTTT 47  
Db 205 CACAGTTTACTGAGATGAATGCGCTGATCTACTATTGAGACTTT 159

RESULT 8  
US-09-949-016-17061  
; Sequence 17061, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17061  
; LENGTH: 96922  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17061

Query Match 51.1%; Score 24; DB 4; Length 96922;  
Best Local Similarity 75.0%; Pred. No. 58;  
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 8 CTTGATTAGAAATAASTTTGATCACCATTTCGAATTTT 47  
Db 34578 CATATTTAGAAATAAATAATGATCCGATATTTAATTTT 34617

RESULT 9  
US-09-949-016-52894  
; Sequence 52894, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20

Query Match 51.5%; Score 24.2; DB 4; Length 993;  
Best Local Similarity 68.1%; Pred. No. 25;  
Matches 32; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

;; PRIOR APPLICATION NUMBER: 60/237,768  
;; PRIOR FILING DATE: 2000-10-03  
;; PRIOR APPLICATION NUMBER: 60/231,498  
;; PRIOR FILING DATE: 2000-09-08  
;; NUMBER OF SEQ ID NOS: 207012  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO: 52894  
;; LENGTH: 601  
;; TYPE: DNA  
;; ORGANISM: Human  
US-09-949-016-52894

Query Match 50.2%; Score 23.6; DB 4; Length 601;  
Best Local Similarity 76.3%; Pred. No. 38;  
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 9 TTGATTAGAAATAAATTGATCACCATTTCATTT 46  
DB 204 TTGATTATATAATAATTTCATATCATTTTCTATT 241

RESULT 10  
US-09-949-016-13271/C  
; Sequence 13271, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 13271  
; LENGTH: 90876  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(90876)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13271

Query Match 50.2%; Score 23.6; DB 4; Length 90876;  
Best Local Similarity 76.3%; Pred. No. 79;  
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 9 TTGATTAGAAATAAATTGATCACCATTTCATTT 46  
DB 24298 TTGATTATATAATAATTTCATATCATTTTCTATT 24261

RESULT 11  
US-09-949-016-185469  
; Sequence 185469, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768

;; PRIOR FILING DATE: 2000-10-03  
;; PRIOR APPLICATION NUMBER: 60/231,498  
;; PRIOR FILING DATE: 2000-09-08  
;; NUMBER OF SEQ ID NOS: 207012  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO: 185469  
;; LENGTH: 601  
;; TYPE: DNA  
;; ORGANISM: Human  
US-09-949-016-185469

Query Match 49.4%; Score 23.2; DB 4; Length 601;  
Best Local Similarity 73.7%; Pred. No. 53;  
Matches 28; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 5 AGGCTTGATTGAATAAATTGATCACCATTTCAA 42  
DB 148 AGGATTCAGTAGAATAAAGTTTATTTTCGTTTCCAA 185

RESULT 12  
US-09-949-016-185470  
; Sequence 185470, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 185470  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-185470

Query Match 49.4%; Score 23.2; DB 4; Length 601;  
Best Local Similarity 73.7%; Pred. No. 53;  
Matches 28; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 5 AGGCTTGATTGAATAAATTGATCACCATTTCAA 42  
DB 140 AGGATTCAGTAGAATAAAGTTTATTTTCGTTTCCAA 177

RESULT 13  
US-09-949-016-1358  
; Sequence 1358, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1358  
LENGTH: 1978  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-1358

Query Match 49.4%; Score 23.2; DB 4; Length 1978;  
Best Local Similarity 83.3%; Pred. No. 63;  
Matches 25; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 10 TGATTGAAATTAATTTGATCACAATTTT 39  
DB 1939 TGACTTGAATTAACCTTGACACAAATTTT 1968

RESULT 14  
US-09-248-796A-1897  
Sequence 1897, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstein et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 1897  
LENGTH: 2064  
TYPE: DNA  
ORGANISM: Candida albicans  
US-09-248-796A-1897

Query Match 49.4%; Score 23.2; DB 4; Length 2064;  
Best Local Similarity 73.7%; Pred. No. 63;  
Matches 28; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 9 TTGATTGAAATTAATTTGATCACAATTTTCAATTT 46  
DB 1609 TTGATTGACATGACTTTATCTACATTTAAACTT 1646

RESULT 15  
US-09-949-016-13100  
Sequence 13100, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13100  
LENGTH: 24538  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-13100

Query Match 49.4%; Score 23.2; DB 4; Length 24538;  
Best Local Similarity 83.3%; Pred. No. 90;  
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Matches 25; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 10 TGATTGAAATTAATTTGATCACAATTTT 39  
DB 22499 TGACTTGAATTAACCTTGACACAAATTTT 22528

RESULT 16  
US-09-949-016-17049  
Sequence 17049, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 17049  
LENGTH: 44971  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...((44971))  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-17049

Query Match 49.4%; Score 23.2; DB 4; Length 44971;  
Best Local Similarity 73.7%; Pred. No. 98;  
Matches 28; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 5 AGCGTTGATTGAAATTAATTTGATCACAATTTTCAA 42  
DB 31128 AGGATTGAGTGAATTAAGTTTATTCGTTTCAA 31165

RESULT 17  
US-09-949-016-12288/c  
Sequence 12288, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12288  
LENGTH: 76164  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-12288

Query Match 48.9%; Score 23; DB 4; Length 76164;  
Best Local Similarity 74.4%; Pred. No. 1.2e+02;  
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY		9	TTCATTGAAATAAATTTTGATCACCATTTCGAATTTT	47
Db		40627	TTCACTAAAAATTAAAATTAAACATTTTAAAACTTT	40589

```

RESULT 18
US-09-949-016-14005/c
: Sequence 14005, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CL001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 14005
: LENGTH: 76165
: TYPE: DNA
: ORGANISM: Human
: US-09-949-016-14005

```

	Query Match	48.9%;	Score 23;	DB 4;	Length 76165;
	Best Local Similarity	74.4%;	Pred. No. 1.2e+02;		
	Matches	29;	Conservative	0;	Mismatches 10; Indels 0; Gaps 0;
OY	9	TTGATAGAAATAAATTGGATCCACCATTTTCAAAATTTT	47		
Db	40627	TTTCATTAATAATTAATAATTAATTAACATTTTAAAACTTT	40589		

```

RESULT 19
US-09-360-186-1/c
; Sequence 1, Application US/09360186
; Patent No. 6262343
; GENERAL INFORMATION:
; APPLICANT: Stasikawicz, et al.
; TITLE OF INVENTION: Bc2 Resistance Gene
; FILE REFERENCE: 50687
; CURRENT APPLICATION NUMBER: US/09360, 186
; CURRENT FILING DATE: 1999-07-23
; EARLIER APPLICATION NUMBER: 60/093, 957
; EARLIER FILING DATE: 1998-07-23
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 31491
; TYPE: DNA
; ORGANISM: Capsicum annuum
; US-09-360-186-1

```

Query Match	48.5%	Score 22.8;	DB 3;	length 31491;
Best Local Similarity	71.4%;	Pred. No. 1.3e+02;		
Matches 30;	Conservative 0;	Mismatches 12;	Indels 0;	Gaps 0;

RESULT 20  
US-09-864-680A-1/c  
; Sequence 1, Application US/09864680A  
; Patent No. 6762285

```

1  GENERAL INFORMATION:
2  APPLICANT: Staekawicz, Brian J
3  APPLICANT: Dahlbeck, Douglas
4  APPLICANT: Tai, Thomas H
5  TITLE OF INVENTION: B92 RESISTANCE GENE
6  FILE REFERENCE: 4,225/0,23,4021 (5830-4A)
7  CURRENT APPLICATION NUMBER: US/09/864,680A
8  CURRENT FILING DATE: 2001-05-24
9  PRIOR APPLICATION NUMBER: US 09/360,186
10 PRIOR FILING DATE: 1999-07-23
11 PRIOR APPLICATION NUMBER: US 60/093,957
12 PRIOR FILING DATE: 1998-07-23
13 NUMBER OF SEQ ID NOS: 14
14
15 SOFTWARE: PatentIn version 3.1
16
17 SEQ ID NO 1
18
19 LENGTH: 31491
20 TYPE: DNA
21
22 ORGANISM: Capsicum annuum
23
24 US-09-864-680A-1

```

Query	3	ACAGCGTGTGATTGAAATTAATTTGATCACCATTTCGAAT	44
Db	27659	AGAGCATGAATTGAAATTAATTTGTCACCACTTGTAAT	27618
Matches	30;	Conservative	0;
		Mismatches	12;
		Indels	0;
		Gaps	0;

```

RESULT 21
US-08-956-171E-1138
Sequence 1138, Application US/08956171E
Patent No. 6593114
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
Gil H. Choi
Patrick S. Dillon
Craig A. Rosen
Steven C. Barash
Michael R. Fannon
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO. 1138:
SEQUENCE CHARACTERISTICS:
LENGTH: 459 base pairs

```



TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1138:  
US-08-956-171E-1138

Query Match  
Best Local Similarity 48.1%; Score 22.6; DB 4; Length 459;  
Matches 31; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

3 ACAGCGTTGATTGAAATTAATTTGATCACCATTTCGAATTTT 47  
317 AAAAAGTTTGTCACAAATTAATCTTATCTCAATTTTAAATAT 361

RESULT 22  
US-08-781-986A-1138  
Sequence 1138, Application US/08781986A  
Patent No. 6737248  
GENERAL INFORMATION:  
APPLICANT: Charles Kunsch  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5255  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/781,986A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benson, Bob  
REGISTRATION NUMBER: 30,446  
REFERENCE/DOCKET NUMBER: PB248PP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 1138:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 459 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-781-986A-1138

Query Match  
Best Local Similarity 48.1%; Score 22.6; DB 4; Length 459;  
Matches 31; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

3 ACAGCGTTGATTGAAATTAATTTGATCACCATTTCGAATTTT 47  
317 AAAAAGTTTGTCACAAATTAATCTTATCTCAATTTTAAATAT 361

RESULT 23  
US-09-949-016-141510  
Sequence 141510, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 141510  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-141510

Query Match  
Best Local Similarity 48.1%; Score 22.6; DB 4; Length 601;  
Matches 28; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

8 CTGATTGAAATTAATTAATTTGATCACCATTTCGAATTT 46  
337 CTGATTAATTAATTAATTAATTAATTAATTAATTAATTT 375

RESULT 24  
US-08-936-165A-55/C  
Sequence 55, Application US/08936165A  
Patent No. 6348582  
GENERAL INFORMATION:  
APPLICANT: Black, Michael  
APPLICANT: Burnham, Martin  
APPLICANT: Hodgson, John  
APPLICANT: Knowles, David  
APPLICANT: Lonetto, Michael  
APPLICANT: Nicholas, Richard  
APPLICANT: Pratt, Julie  
APPLICANT: Reichard, Richard  
APPLICANT: Rosenberg, Martin  
APPLICANT: Ward, Judith  
TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,  
NUMBER OF SEQUENCES: 534  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/936,165A  
FILING DATE: 24-SEP-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/027,032  
FILING DATE: 24-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Glumel, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P50549  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 55:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 1209 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: genomic DNA  
 US-08-936-165A-55

Query Match 48.1%; Score 22.6; DB 3; Length 1209;  
 Best Local Similarity 68.9%; Pred. No. 94;  
 Matches 31; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 3 ACAGCCTTGATTGAAATTAATTAATTTGATCACCATTTCATAATTTT 47  
 Db 439 AAAAATTGTCACAAATTAATCTTATCTCAATTTTAAATATT 395

RESULT 25  
 US-09-949-016-16430

; Sequence 16430, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: fastSeq for Windows Version 4.0  
 ; SEQ ID NO 16430  
 ; LENGTH: 57726  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-949-016-16430

Query Match 48.1%; Score 22.6; DB 4; Length 57726;  
 Best Local Similarity 68.9%; Pred. No. 1.6e+02;  
 Matches 31; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CCACAGCCTTGATTGAAATTAATTAATTTGATCACCATTTCATAATT 45  
 Db 48861 CCATTAAGGTAGATTAACATTAATTTTCCGTATTAATTT 48905

Search completed: June 2, 2005, 10:22:43  
 Job time : 136 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: June 2, 2005, 08:49:54 ; Search time 517 Seconds  
(without alignments)  
558.857 Million cell updates/sec

Title: US-10-349-143-1

Perfect score: 47  
Sequence: 1 ccacagcgtcgttagtaaat.....gattcaccatttcacatttt 47

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5706582 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46.6	99.1	47	US-10-349-143-1	Sequence 1, Appl1
2	46.6	99.1	47	US-10-349-143-1	Sequence 1, Appl1
3	46.6	99.1	47	US-10-349-143-1	Sequence 1, Appl1
4	24.8	52.1	18	US-10-468-582-1	Sequence 1, Appl1
5	24.8	52.1	18	US-10-719-993-29983	Sequence 29983, A
6	24.4	51.9	18	US-10-719-993-6643	Sequence 6643, Ap
7	24.4	51.9	9	US-09-796-692-4373	Sequence 4373, Ap
8	24.4	51.9	9	US-09-796-692-4373	Sequence 4373, Ap
9	24.4	51.9	14	US-10-040-862-4373	Sequence 4373, Ap
10	24.4	51.9	14	US-10-040-862-4373	Sequence 4373, Ap
11	24.4	51.9	267	US-10-057-475B-4373	Sequence 4373, Ap
12	24.4	51.9	267	US-10-057-475B-4373	Sequence 4373, Ap

12	24.4	51.9	267	US-10-154-884B-4373	Sequence 4373, Ap
13	24.4	51.9	267	US-10-154-884B-4373	Sequence 4373, Ap
14	24.4	51.9	267	US-10-764-324-4373	Sequence 9383, Ap
15	24.4	51.9	267	US-10-764-324-4373	Sequence 9383, Ap
16	24.4	51.9	589	US-10-027-633-240198	Sequence 240198, A
17	24.4	51.9	589	US-10-027-633-240198	Sequence 240198, A
18	24.4	51.9	589	US-10-027-633-240198	Sequence 240198, A
19	24.4	51.9	589	US-10-027-633-240198	Sequence 240198, A
20	24.4	51.9	589	US-10-027-633-240198	Sequence 240198, A
21	24.4	51.9	589	US-10-027-633-240198	Sequence 240198, A
22	24.2	51.5	2199	US-10-437-963-51548	Sequence 51548, A
23	24.2	51.5	5173	US-10-311-455-924	Sequence 924, Ap
24	23.8	50.6	728	US-10-767-701-2646	Sequence 2646, Ap
25	23.8	50.6	1102	US-10-425-114-15544	Sequence 15544, A
26	23.8	50.6	1105	US-10-425-115-83159	Sequence 83159, A
27	23.6	50.2	27681	US-09-764-869-1997	Sequence 1997, Ap
28	23.6	50.2	27681	US-09-764-869-1997	Sequence 1997, Ap
29	23.6	50.2	27681	US-10-091-504-1997	Sequence 1997, Ap
30	23.6	50.2	27681	US-10-091-504-1997	Sequence 1997, Ap
31	23.6	50.2	27681	US-10-227-577-1997	Sequence 1997, Ap
32	23.6	50.2	27681	US-10-227-577-1997	Sequence 1997, Ap
33	23.4	49.8	366	US-10-425-115-182292	Sequence 182292, A
34	23.4	49.8	562	US-10-027-633-45686	Sequence 45686, A
35	23.4	49.8	562	US-10-027-633-45686	Sequence 45686, A
36	23.4	49.8	562	US-10-027-633-45686	Sequence 45686, A
37	23.4	49.8	562	US-10-027-633-45686	Sequence 45686, A
38	23.4	49.8	562	US-10-027-633-45686	Sequence 45686, A
39	23.4	49.8	562	US-10-027-633-45686	Sequence 45686, A
40	23.4	49.8	721	US-10-027-633-155815	Sequence 155815, A
41	23.4	49.8	721	US-10-027-633-155815	Sequence 155815, A
42	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
43	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
44	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
45	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
46	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
47	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
48	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
49	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
50	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
51	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
52	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
53	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
54	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
55	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
56	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
57	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
58	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
59	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
60	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
61	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
62	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
63	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
64	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
65	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
66	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
67	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
68	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
69	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
70	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
71	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
72	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
73	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
74	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
75	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
76	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
77	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
78	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
79	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
80	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
81	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
82	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
83	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap
84	23.2	49.4	370	US-10-424-559-1809	Sequence 1809, Ap

```
C 85 22.6 48.1 611 17 US-10-027-632-203374 Sequence 203374,
C 86 22.6 48.1 611 17 US-10-027-632-203375 Sequence 203375,
C 87 22.6 48.1 701 18 US-10-653-047-6311 Sequence 6311, Ap
C 88 22.6 48.1 730 13 US-10-027-632-24249 Sequence 24249, A
C 89 22.6 48.1 730 17 US-10-027-632-24249 Sequence 24249, A
C 90 22.6 48.1 792 13 US-10-027-632-127002 Sequence 127002,
C 91 22.6 48.1 792 13 US-10-027-632-127003 Sequence 127003,
C 92 22.6 48.1 792 13 US-10-027-632-165053 Sequence 165053,
C 93 22.6 48.1 792 17 US-10-027-632-127002 Sequence 127002,
C 94 22.6 48.1 792 17 US-10-027-632-127003 Sequence 127003,
C 95 22.6 48.1 792 17 US-10-027-632-165053 Sequence 165053,
C 96 22.6 48.1 1209 9 US-09-939-980-55 Sequence 55, Appl
C 97 22.6 48.1 2319 18 US-10-437-963-87185 Sequence 87185, A
C 98 22.6 48.1 2526 17 US-10-425-114-31597 Sequence 31597, A
C 99 22.6 48.1 73725 17 US-10-052-482-196 Sequence 196, App
C 100 22.6 48.1 322101 16 US-10-060-902-1 Sequence 1, Appli
```

## ALIGNMENTS

```
RESULT 1
US-10-349-143-1
; Sequence 1, Application US/10349143
; Publication No. US20040005584A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marcia
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSET.020CPI
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: US/10/349,143
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US/09/422,978
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/298,850
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/109,732
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-23
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/082,614
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 1
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 99-109-224 : polymorphic base G or C
US-10-349-143-1

Query Match 99.1%; Score 46.6; DB 17; Length 47;
Best Local Similarity 100.0%; Pred. No. 8.9e-06;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CCACAGGCTTGATGAATAAATTAATTTGATTCACCAATTTCAATTTT 47
Db 1 CCACAGGCTTGATGAATAAATTAATTTGATTCACCAATTTCAATTTT 47

RESULT 2
US-09-790-289-1
; Sequence 1, Application US/09790289
; Publication No. US20030165826A1
; GENERAL INFORMATION:
; APPLICANT: Caroline Barry
; APPLICANT: Ilya Chumakov
; TITLE OF INVENTION: PG-3 and Biallelic Markers Thereof
; FILE REFERENCE: 68 US3 REG
; CURRENT APPLICATION NUMBER: US/09/790,289
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 5
```

```
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 240825
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..2000
; OTHER INFORMATION: 5'regulatory region
; NAME/KEY: exon
; LOCATION: 2001..2079
; OTHER INFORMATION: exon A
; NAME/KEY: exon
; LOCATION: 4627..4718
; OTHER INFORMATION: exon B
; NAME/KEY: exon
; LOCATION: 10115..10233
; OTHER INFORMATION: exon C
; NAME/KEY: exon
; LOCATION: 26810..26897
; OTHER INFORMATION: exon D
; NAME/KEY: exon
; LOCATION: 31357..31471
; OTHER INFORMATION: exon E
; NAME/KEY: exon
; LOCATION: 34261..34404
; OTHER INFORMATION: exon F
; NAME/KEY: exon
; LOCATION: 37377..37466
; OTHER INFORMATION: exon S
; NAME/KEY: exon
; LOCATION: 39704..40858
; OTHER INFORMATION: exon T
; NAME/KEY: exon
; LOCATION: 50436..50545
; OTHER INFORMATION: exon G
; NAME/KEY: exon
; LOCATION: 72881..72918
; OTHER INFORMATION: exon H
; NAME/KEY: exon
; LOCATION: 75989..76151
; OTHER INFORMATION: exon I
; NAME/KEY: exon
; LOCATION: 95111..95188
; OTHER INFORMATION: exon J
; NAME/KEY: exon
; LOCATION: 216015..216252
; OTHER INFORMATION: exon K
; NAME/KEY: exon
; LOCATION: 237526..238825
; OTHER INFORMATION: exon L
; NAME/KEY: misc.feature
; LOCATION: 238826..240825
; OTHER INFORMATION: 3'regulatory region
; NAME/KEY: allele
; LOCATION: 1999
; OTHER INFORMATION: 5-390-177 : polymorphic base G or C
; NAME/KEY: allele
; LOCATION: 4601
; OTHER INFORMATION: 5-391-43 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 10228
; OTHER INFORMATION: 5-392-222 : polymorphic base G or T
; NAME/KEY: allele
; LOCATION: 10286
; OTHER INFORMATION: 5-392-280 : polymorphic base G or T
; NAME/KEY: allele
; LOCATION: 10370
; OTHER INFORMATION: 5-392-364 : insertion of G
; NAME/KEY: allele
; LOCATION: 39944
; OTHER INFORMATION: 4-58-318 : polymorphic base G or T
; NAME/KEY: allele
```

LOCATION: 39973  
OTHER INFORMATION: 4-58-289 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 41385  
OTHER INFORMATION: 4-54-199 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 41404  
OTHER INFORMATION: 4-54-180 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 42232  
OTHER INFORMATION: 4-51-312 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 67475  
OTHER INFORMATION: 99-86-266 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 69521  
OTHER INFORMATION: 4-88-107 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 72838  
OTHER INFORMATION: 5-397-141 : polymorphic base G or T  
NAME/KEY: allele  
LOCATION: 76060  
OTHER INFORMATION: 5-398-203 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 81253  
OTHER INFORMATION: 99-12738-248 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 83921  
OTHER INFORMATION: 99-109-358 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 91917  
OTHER INFORMATION: 99-12749-175 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 95349  
OTHER INFORMATION: 4-21-154 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 95511  
OTHER INFORMATION: 4-21-317 : polymorphic base G or T  
NAME/KEY: allele  
LOCATION: 96190  
OTHER INFORMATION: 4-23-326 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 97294  
OTHER INFORMATION: 99-12753-34 : polymorphic base A or T  
NAME/KEY: allele  
LOCATION: 98024  
OTHER INFORMATION: 5-364-252 : polymorphic base G or T  
NAME/KEY: allele  
LOCATION: 98914  
OTHER INFORMATION: 99-12755-280 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 98963  
OTHER INFORMATION: 99-12755-329 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 103593  
OTHER INFORMATION: 4-87-212 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 104398  
OTHER INFORMATION: 99-12757-118 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 106373  
OTHER INFORMATION: 99-12758-102 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 106407  
OTHER INFORMATION: 99-12758-136 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 108315  
OTHER INFORMATION: 4-105-98 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 108327  
OTHER INFORMATION: 4-105-86 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 108472

OTHER INFORMATION: 4-45-49 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 109196  
OTHER INFORMATION: 4-44-277 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 114604  
OTHER INFORMATION: 4-86-60 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 115716  
OTHER INFORMATION: 4-84-334 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 122083  
OTHER INFORMATION: 99-78-321 : polymorphic base A or T  
NAME/KEY: allele  
LOCATION: 123124  
OTHER INFORMATION: 99-12767-36 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 123231  
OTHER INFORMATION: 99-12767-143 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 123277  
OTHER INFORMATION: 99-12767-189 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 123468  
OTHER INFORMATION: 99-12767-380 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 126738  
OTHER INFORMATION: 4-80-328 : polymorphic base C or T  
NAME/KEY: allele  
LOCATION: 128210  
OTHER INFORMATION: 4-36-384 : polymorphic base G or C  
NAME/KEY: allele  
LOCATION: 128330  
OTHER INFORMATION: 4-36-264 : polymorphic base A or G  
NAME/KEY: allele  
LOCATION: 128333  
OTHER INFORMATION: 4-36-261 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 128594  
OTHER INFORMATION: 4-35-333 : polymorphic base A or C  
NAME/KEY: allele  
LOCATION: 128687  
OTHER INFORMATION: 4-35-240 : polymorphic base G or C  
NAME/KEY: allele

Query Match 99.1%; Score 46.6; DB 10; Length 240825;  
Best Local Similarity 97.9%; Pred. No. 8.4e-05;  
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACAGCCTTGATTAATAASTTGTGATCCCATTTCAAAATTT 47  
DB 83764 CCACAGCCTTGATTAATAATAAGTTGATCACCATTTCAAATTT 83810

RESULT 3  
US-10-468-582-1  
Sequence 1, Application US/10468582  
Publication No. US20040163137A1  
GENERAL INFORMATION:  
APPLICANT: Genset  
TITLE OF INVENTION: PG-3 and bisallelic markers thereof  
FILE REFERENCE: 68. W02  
CURRENT APPLICATION NUMBER: US/10/468,582  
CURRENT FILING DATE: 2003-08-20  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: Patent.pm  
SEQ ID NO 1  
LENGTH: 240825  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1..2000

OTHER INFORMATION: 5'regulatory region  
FEATURE:  
NAME/KEY: exon  
LOCATION: 2001..2079  
OTHER INFORMATION: exon A  
FEATURE:  
NAME/KEY: exon  
LOCATION: 4627..4718  
OTHER INFORMATION: exon B  
FEATURE:  
NAME/KEY: exon  
LOCATION: 10115..10233  
OTHER INFORMATION: exon C  
FEATURE:  
NAME/KEY: exon  
LOCATION: 26810..26897  
OTHER INFORMATION: exon D  
FEATURE:  
NAME/KEY: exon  
LOCATION: 31357..31471  
OTHER INFORMATION: exon E  
FEATURE:  
NAME/KEY: exon  
LOCATION: 34261..34404  
OTHER INFORMATION: exon F  
FEATURE:  
NAME/KEY: exon  
LOCATION: 37377..37466  
OTHER INFORMATION: exon S  
FEATURE:  
NAME/KEY: exon  
LOCATION: 39704..40858  
OTHER INFORMATION: exon T  
FEATURE:  
NAME/KEY: exon  
LOCATION: 50436..50545  
OTHER INFORMATION: exon G  
FEATURE:  
NAME/KEY: exon  
LOCATION: 72881..72918  
OTHER INFORMATION: exon H  
FEATURE:  
NAME/KEY: exon  
LOCATION: 75989..76151  
OTHER INFORMATION: exon I  
FEATURE:  
NAME/KEY: exon  
LOCATION: 95111..95188  
OTHER INFORMATION: exon J  
FEATURE:  
NAME/KEY: exon  
LOCATION: 216015..216252  
OTHER INFORMATION: exon K  
FEATURE:  
NAME/KEY: exon  
LOCATION: 237526..238825  
OTHER INFORMATION: exon L  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 238826..240825  
OTHER INFORMATION: 3'regulatory region  
FEATURE:  
NAME/KEY: allele  
LOCATION: 1999  
OTHER INFORMATION: 5-390-177 : polymorphic base G or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 4601  
OTHER INFORMATION: 5-391-43 : polymorphic base A or G  
FEATURE:  
NAME/KEY: allele  
LOCATION: 10228  
OTHER INFORMATION: 5-392-222 : polymorphic base G or T

FEATURE:  
NAME/KEY: allele  
LOCATION: 10286  
OTHER INFORMATION: 5-392-280 : polymorphic base G or T  
FEATURE:  
NAME/KEY: allele  
LOCATION: 10370  
OTHER INFORMATION: 5-392-364 : insertion of G  
FEATURE:  
NAME/KEY: allele  
LOCATION: 39944  
OTHER INFORMATION: 4-58-318 : polymorphic base G or T  
FEATURE:  
NAME/KEY: allele  
LOCATION: 39973  
OTHER INFORMATION: 4-58-289 : polymorphic base G or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 41385  
OTHER INFORMATION: 4-54-199 : polymorphic base A or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 41404  
OTHER INFORMATION: 4-54-180 : polymorphic base A or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 42232  
OTHER INFORMATION: 4-51-312 : polymorphic base G or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 67475  
OTHER INFORMATION: 99-86-266 : polymorphic base A or G  
FEATURE:  
NAME/KEY: allele  
LOCATION: 69521  
OTHER INFORMATION: 4-88-107 : polymorphic base A or G  
FEATURE:  
NAME/KEY: allele  
LOCATION: 72838  
OTHER INFORMATION: 5-397-141 : polymorphic base G or T  
FEATURE:  
NAME/KEY: allele  
LOCATION: 76060  
OTHER INFORMATION: 5-398-203 : polymorphic base A or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 81253  
OTHER INFORMATION: 99-12738-248 : polymorphic base A or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 83921  
OTHER INFORMATION: 99-109-358 : polymorphic base A or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 91917  
OTHER INFORMATION: 99-12749-175 : polymorphic base C or T  
FEATURE:  
NAME/KEY: allele  
LOCATION: 95349  
OTHER INFORMATION: 4-21-154 : polymorphic base C or T  
FEATURE:  
NAME/KEY: allele  
LOCATION: 95511  
OTHER INFORMATION: 4-21-317 : polymorphic base G or T  
FEATURE:  
NAME/KEY: allele  
LOCATION: 96190  
OTHER INFORMATION: 4-23-326 : polymorphic base A or G  
FEATURE:  
NAME/KEY: allele  
LOCATION: 97294  
OTHER INFORMATION: 99-12753-34 : polymorphic base A or T  
FEATURE:

NAME/KEY: allele  
LOCATION: 98024  
OTHER INFORMATION: 5-364-252 : polymorphic base G or T  
FEATURE:  
NAME/KEY: allele  
LOCATION: 98914  
OTHER INFORMATION: 99-12755-280 : polymorphic base A or G  
FEATURE:  
NAME/KEY: allele  
LOCATION: 98963  
OTHER INFORMATION: 99-12755-329 : polymorphic base A or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 103593  
OTHER INFORMATION: 4-87-212 : polymorphic base A or G  
FEATURE:  
NAME/KEY: allele  
LOCATION: 104398  
OTHER INFORMATION: 99-12757-318 : polymorphic base C or T  
FEATURE:  
NAME/KEY: allele  
LOCATION: 106373  
OTHER INFORMATION: 99-12758-102 : polymorphic base A or G  
FEATURE:  
NAME/KEY: allele  
LOCATION: 106407  
OTHER INFORMATION: 99-12758-136 : polymorphic base C or T  
FEATURE:  
NAME/KEY: allele  
LOCATION: 108315  
OTHER INFORMATION: 4-105-98 : polymorphic base A or G  
FEATURE:  
NAME/KEY: allele  
LOCATION: 108327  
OTHER INFORMATION: 4-105-86 : polymorphic base A or G  
FEATURE:  
NAME/KEY: allele

Query Match 99.1%; Score 46.6; DB 18; Length 240825;  
Best Local Similarity 97.9%; Pred. No. 8.4e-05;  
Matches 46; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACGGCTGATTAAGTAATTAATTTGATCCCAATTTCAATTTT 47  
DB 83764 CCACGGCTGATTAAGTAATTAATTTGATCCCAATTTCAATTTT 83810

## RESULT 4

US-10-719-993-29983  
Sequence 29983, Application US/10719993  
Publication No. US20040265849A1  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: CL001496  
CURRENT APPLICATION NUMBER: US/10/719,993  
NUMBER OF SEQ ID NOS: 55342  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 29983  
LENGTH: 201  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-719-993-29983

Query Match 52.8%; Score 24.8; DB 18; Length 201;  
Best Local Similarity 76.3%; Pred. No. 1.9e+02;  
Matches 29; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 10 TGATTGAATAAATTTGATCCCAATTTCAATTTT 47  
DB 71 TTATAGAAATTAACCTTAATTAATTCCTTTTAAATTT 108

RESULT 5  
US-10-719-993-6843  
Sequence 6843, Application US/10719993  
Publication No. US20040265849A1  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: CL001496  
CURRENT APPLICATION NUMBER: US/10/719,993  
NUMBER OF SEQ ID NOS: 55342  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6843  
LENGTH: 108111  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-719-993-6843

Query Match 52.8%; Score 24.8; DB 18; Length 108111;  
Best Local Similarity 76.3%; Pred. No. 9.9e+02;  
Matches 29; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 10 TGATTGAATAAATTTGATCCCAATTTCAATTTT 47  
DB 11145 TTATAGAAATTAACCTTAATTAATTCCTTTTAAATTT 11182

## RESULT 6

US-09-796-692-4373/C  
Sequence 4373, Application US/09796692  
Publication No. US20020198362A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Mannion, Jane  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
FILE REFERENCE: 2077.001200  
CURRENT APPLICATION NUMBER: US/09/796,692  
CURRENT FILING DATE: 2001-03-01  
PRIOR APPLICATION NUMBER: 60/186,126  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: 60/200,545  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 60/200,303  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,779  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,999  
PRIOR FILING DATE: 2000-05-01  
PRIOR APPLICATION NUMBER: 60/202,084  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: 60/206,201  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: 60/218,950  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: 60/222,903  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: 60/223,416  
PRIOR FILING DATE: 2000-08-04  
PRIOR APPLICATION NUMBER: 60/223,378  
PRIOR FILING DATE: 2000-08-07  
NUMBER OF SEQ ID NOS: 9597  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4373  
LENGTH: 267  
TYPE: DNA  
ORGANISM: Homo sapiens

US-09-796-692-4373

Query Match 51.9%; Score 24.4; DB 9; Length 267;  
Best Local Similarity 70.5%; Pred. No. 2.8e+02;  
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

Qy 4 CAGCGTTGATTGAAATAAATTTGATCACCATTTCATTTT 47  
Db 196 CATGCTTGCTTTTATATCAACTGTATATCACCATTATTTATTTT 153

RESULT 7

US-09-796-692-9383/c

; Sequence 9383, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09/796,692  
; PRIOR FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; NUMBER OF SEQ ID NOS: 9597  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 9383  
; LENGTH: 267  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-796-692-9383

US-09-796-692-9383

Query Match 51.9%; Score 24.4; DB 9; Length 267;  
Best Local Similarity 70.5%; Pred. No. 2.8e+02;  
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

Qy 4 CAGCGTTGATTGAAATAAATTTGATCACCATTTCATTTT 47  
Db 196 CATGCTTGCTTTTATATCAACTGTATATCACCATTATTTATTTT 153

RESULT 8

US-10-040-862-4373/c

; Sequence 4373, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane

; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013520US  
; CURRENT APPLICATION NUMBER: US/10/040,862  
; CURRENT FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: US 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: US 09/796,692  
; PRIOR FILING DATE: 2001-03-01  
; NUMBER OF SEQ ID NOS: 10467  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4373  
; LENGTH: 267  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-040-862-4373

Query Match 51.9%; Score 24.4; DB 14; Length 267;  
Best Local Similarity 70.5%; Pred. No. 2.8e+02;  
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

Qy 4 CAGCGTTGATTGAAATAAATTTGATCACCATTTCATTTT 47  
Db 196 CATGCTTGCTTTTATATCAACTGTATATCACCATTATTTATTTT 153

RESULT 9

US-10-040-862-9383/c

; Sequence 9383, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013520US  
; CURRENT APPLICATION NUMBER: US/10/040,862  
; CURRENT FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28



```

; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 9383
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-040-862-9383

Query Match      51.9%; Score 24.4; DB 14; Length 267;
Best Local Similarity 70.5%; Pred. No. 2.8e+02;
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 4 CAGGCTGATTGAAATAAATTGATCACCATTTCGAATTT 47
DB 196 CATGCTGCTTTTATTCACGTGTATCACCATTTATTAATTT 153

RESULT 10
; Sequence 4373, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aljun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
```

```

; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 4373
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-057-475B-9383

Query Match      51.9%; Score 24.4; DB 17; Length 267;
Best Local Similarity 70.5%; Pred. No. 2.8e+02;
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 4 CAGGCTGATTGAAATAAATTGATCACCATTTCGAATTT 47
DB 196 CATGCTGCTTTTATTCACGTGTATCACCATTTATTAATTT 153

RESULT 11
; Sequence 9383, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aljun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 9383
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-057-475B-9383

Query Match      51.9%; Score 24.4; DB 17; Length 267;
Best Local Similarity 70.5%; Pred. No. 2.8e+02;
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 4 CAGGCTGATTGAAATAAATTGATCACCATTTCGAATTT 47
DB 196 CATGCTGCTTTTATTCACGTGTATCACCATTTATTAATTT 153
```

```
DB      196 CAGCTTGCTTTTATCACTGTATCACCATTATTAATTTT 153

RESULT 12
US-10-154-884B-4373/c
; Sequence 4373, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4373
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-4373

Query Match      51.9%; Score 24.4; DB 17; Length 267;
Best Local Similarity 70.5%; Pred. No. 2.8e+02;
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY      4 CAGCGTTGATTGAAATAASTTTGATCACCATTTCAAATTTT 47
DB      196 CAGCTTGCTTTTATCACTGTATCACCATTATTAATTTT 153

RESULT 13
US-10-154-884B-9383/c
; Sequence 9383, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; PRIOR FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
```

```
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9383
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-9383

Query Match      51.9%; Score 24.4; DB 17; Length 267;
Best Local Similarity 70.5%; Pred. No. 2.8e+02;
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY      4 CAGCGTTGATTGAAATAASTTTGATCACCATTTCAAATTTT 47
DB      196 CAGCTTGCTTTTATCACTGTATCACCATTATTAATTTT 153

RESULT 14
US-10-764-324-4373/c
; Sequence 4373, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
```

SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4373  
; LENGTH: 267  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-764-324-4373

Query Match 51.9%; Score 24.4; DB 18; Length 267;  
Best Local Similarity 70.5%; Pred. No. 2.8e+02;  
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 4 CAGGCTTGATTGAAATAAATTTGATCACCATTTCATTTTCAATTTT 47  
DB 196 CATGCTTGCTTTTTCATCACTGTATCACCATTATTTATTTT 153

RESULT 15  
US-10-764-324-9383/c  
; Sequence 9383, Application US/10764324  
; Publication No. US20040175739A1  
; GENERAL INFORMATION:  
; APPLICANT: Galger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; FILE REFERENCE: 014058-013520US  
; CURRENT APPLICATION NUMBER: US/10/764,324  
; CURRENT FILING DATE: 2004-01-23  
; PRIOR APPLICATION NUMBER: US/10/040,862  
; PRIOR FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; Remaining Prior Application data removed - See file wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 10467  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 9383  
; LENGTH: 267  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-764-324-9383

Query Match 51.9%; Score 24.4; DB 18; Length 267;  
Best Local Similarity 70.5%; Pred. No. 2.8e+02;  
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 4 CAGGCTTGATTGAAATAAATTTGATCACCATTTCATTTTCAATTTT 47  
DB 196 CATGCTTGCTTTTTCATCACTGTATCACCATTATTTATTTT 153

RESULT 16  
US-10-027-632-240198/c  
; Sequence 240198, Application US/10027632  
; Publication No. US20020198371A1

GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 240198  
; LENGTH: 589  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-240198

Query Match 51.9%; Score 24.4; DB 13; Length 589;  
Best Local Similarity 70.5%; Pred. No. 3.4e+02;  
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 4 CAGGCTTGATTGAAATAAATTTGATCACCATTTCATTTTCAATTTT 47  
DB 553 CATGCTTGCTTTTTCATCACTGTATCACCATTATTTATTTT 510

RESULT 17  
US-10-027-632-240198/c  
; Sequence 240199, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 240199  
; LENGTH: 589  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-240199

Query Match 51.9%; Score 24.4; DB 13; Length 589;  
Best Local Similarity 70.5%; Pred. No. 3.4e+02;



PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 240200  
LENGTH: 589  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-240200

Query Match 51.9%; Score 24.4; DB 17; Length 589;  
Best Local Similarity 70.5%; Pred. No. 3.4e+02;  
Matches 31; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 4 CAGGCTTGATTGAAATAAATTGATCAGCATTTCGAATTT 47  
DB 553 CATTGCTTCTTTTTCATCACTGTTATCCACATTATTATTTT 510

RESULT 22  
US-10-437-963-51548  
Sequence 51548, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 51548  
LENGTH: 2199  
TYPE: DNA  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_5392C.1  
US-10-437-963-51548

Query Match 51.5%; Score 24.2; DB 18; Length 2199;  
Best Local Similarity 71.1%; Pred. No. 5.6e+02;  
Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 CCACGGCTTGATTGAAATAAATTGATCAGCATTTCGAATTT 45  
DB 82 CCAAGGCTCAATTGACACAAATTTCAATTAATTCACACT 126

RESULT 23  
US-10-311-455-924/C  
Sequence 924, Application US/10311455  
Publication No. US20030143606A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
APPLICANT: BERLIN, Kurt

TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining  
FILE OF INVENTION: cytosine methylation  
FILE REFERENCE: 5013.1014  
CURRENT APPLICATION NUMBER: US/10/311,455  
CURRENT FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: PCT/EP01/07537  
PRIOR FILING DATE: 2001-07-02  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 2424  
SEQ ID NO 924  
LENGTH: 5173  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 2436, 2554, 2736, 3004, 4504, 4580  
OTHER INFORMATION: n is a or g or c or t  
US-10-311-455-924

Query Match 51.5%; Score 24.2; DB 15; Length 5173;  
Best Local Similarity 71.1%; Pred. No. 7e+02;  
Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 CCACGGCTTGATTGAAATAAATTGATCAGCATTTCGAATTT 45  
DB 3820 CCATATCTTAATTAAATTAATTAATTCATTAATTCCTAATTT 3776

RESULT 24  
US-10-767-701-2646  
Sequence 2646, Application US/10767701  
Publication No. US20040172684A1  
GENERAL INFORMATION:  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53535)B  
CURRENT APPLICATION NUMBER: US/10/767,701  
CURRENT FILING DATE: 2004-01-29  
NUMBER OF SEQ ID NOS: 63128  
SEQ ID NO 2646  
LENGTH: 728  
TYPE: DNA  
ORGANISM: Sorghum bicolor  
FEATURE:  
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS44741\_1  
US-10-767-701-2646

Query Match 50.6%; Score 23.8; DB 18; Length 728;  
Best Local Similarity 75.7%; Pred. No. 5.7e+02;  
Matches 28; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 11 GATTGAAATAAATTGATCAGCATTTCGAATTT 47  
DB 686 GATTAAACAGAGTTATTAAACAATTTCAATTTT 722

RESULT 25  
US-10-425-114-15544  
Sequence 15544, Application US/10425114  
Publication No. US20040034888A1  
GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Zhou, Yihua  
APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; NUMBER OF SEQ ID NOS: 2003-04-28  
; SEQ ID NO 15544  
; LENGTH: 1102  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3060-021-G2\_F11  
US-10-425-114-15544

Query Match 50.6%; Score 23.8; DB 17; Length 1102;  
Best Local Similarity 75.7%; Pred. No. 6.3e+02;  
Matches 28; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 11 GATTGAAATATAASTTTGATCACCATTTCATAATTTT 47  
DB 625 GAGTGATATATAAGTGTCTCATCCATATTACCGTTT 661

Search completed: June 2, 2005, 10:31:30  
Job time : 525 secs